The Annals of Probability
2020, Vol. 48, No. 3, 1286–1316
https://doi.org/10.1214/19-AOP1391

OPERATOR LIMIT OF THE CIRCULAR BETA ENSEMBLE

BY BENEDEK VALKÓ[1] AND BÁLINT VIRÁG[2]

[1]*Department of Mathematics, University of Wisconsin—Madison, valko@math.wisc.edu*
[2]*Department of Mathematics, University of Toronto, balint@math.toronto.edu*

We provide a precise coupling of the finite circular beta ensembles and their limit process via their operator representations. We prove explicit bounds on the distance of the operators and the corresponding point processes. We also prove an estimate on the beta-dependence of the Sine_β process.

1. Introduction. Valkó and Virág (2017) introduced a family of differential operators parametrized by a path $\gamma : [0,1) \to \mathbb{H}$ in the upper half plane and two points on the boundary. When the path is a certain hyperbolic random walk in the Poincaré half-plane model, the operator $\mathtt{Circ}_{\beta,n}$ has eigenvalues given by the points of the circular beta ensemble scaled and lifted periodically to the real line. With the path $\gamma(t) = \mathcal{B}(-\frac{4}{\beta}\log(1-t))$ where $\mathcal{B}$ is standard hyperbolic Brownian motion, the operator $\mathtt{Sine}_\beta$ has eigenvalues given by the Sine_β process, the limit of the circular beta ensembles. (See Theorems 7 and 8.)

The inverses of these operators in a compatible basis are integral operators denoted by $\mathtt{rCirc}_{\beta,n}$ and $\mathtt{rSine}_\beta$, respectively. Our main result is a coupling which gives $\mathtt{Circ}_{\beta,n} \to \mathtt{Sine}_\beta$ with an explicit rate of convergence. (See Section 2 for additional details, and Figure 1 for an illustration of the coupling.)

THEOREM 1. *There is a probability space with a standard hyperbolic Brownian motion $\mathcal{B}$ and an array of stopping times $0 = \tau_{n,n} < \tau_{n,n-1} < \cdots < \tau_{n,0} = \infty$ so that $\mathcal{B}(\tau_{n,\lceil(1-t)n\rceil}), t \in [0,1)$ has the law of the random walk on $\mathbb{H}$ used to generate $\mathtt{Circ}_{\beta,n}$. This provides a coupling of $\mathtt{Sine}_\beta$ and the sequence of operators $\mathtt{Circ}_{\beta,n}$.*

There exists an a.s. finite positive random variable N so that in this coupling

$$\|\mathtt{rSine}_\beta - \mathtt{rCirc}_{\beta,n}\|_{\mathrm{HS}}^2 \le \frac{\log^6 n}{n} \tag{1}$$

a.s. in the Hilbert–Schmidt norm for all $n \ge N$.

As a corollary, we get new results about the rate of convergence of the eigenvalue processes. Let $\lambda_k, k \in \mathbb{Z}$ be the ordered sequence of eigenvalues of $\mathtt{Sine}_\beta$ with $\lambda_0 < 0 \le \lambda_1$, the sequence $\lambda_{k,n}, k \in \mathbb{Z}$ is defined analogously for $\mathtt{Circ}_{\beta,n}$.

COROLLARY 2. *In the coupling of Theorem 1, we have a.s.,*

$$\sum_{k\in\mathbb{Z}} (\lambda_k^{-1} - \lambda_{n,k}^{-1})^2 \le \frac{\log^6 n}{n} \tag{2}$$

for all $n \ge N$. (Here, N is the finite random variable from Theorem 1.) Moreover, as $n \to \infty$ we have a.s.

$$\max_{|k| \le \frac{n^{1/4}}{\log^2 n}} |\lambda_k - \lambda_{k,n}| \to 0. \tag{3}$$

Received February 2018; revised July 2019.
MSC2010 subject classifications. Primary 60B20, 47B80; secondary 47E05.
Key words and phrases. Random matrices, random operators, coupling.

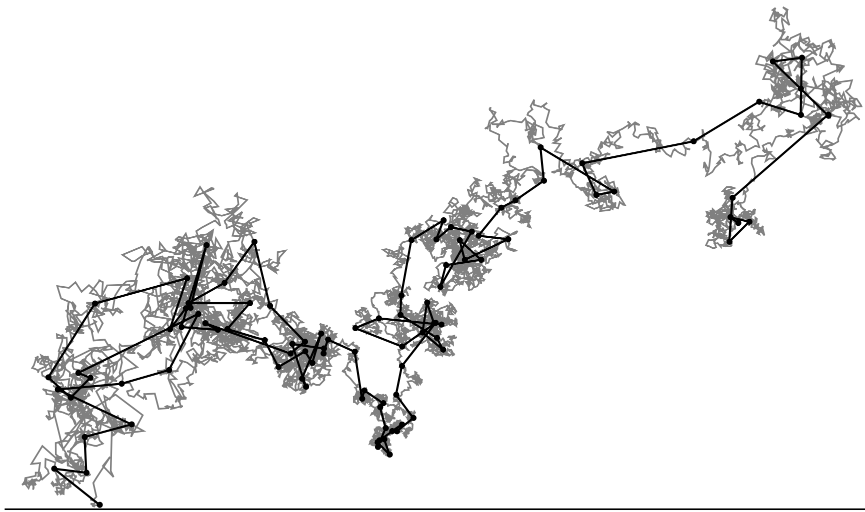

FIG. 1. *Simulation of hyperbolic Brownian motion and a coupled random walk.*

For all $\varepsilon > 0$, *there is a random* N_ε *so that for* $n \geq N_\varepsilon$ *and* $|k| \leq n^{1/2-\varepsilon}$ *we have a.s.*

$$|\lambda_k - \lambda_{k,n}| \leq \frac{1+k^2}{n^{1/2-\varepsilon}}. \tag{4}$$

This provides the best known coupling of the circular β-ensemble to the Sine_β process, even for $\beta = 2$, when both processes are determinantal with explicitly given kernels. For $\beta = 2$, the bound (4) improves on a coupling given in Maples, Najnudel and Nikeghbali (2019) in which the inequality holds with the exponent $1/3$ instead of $1/2$, for $|k| \leq n^{1/4}$.

Using the techniques introduced in our proof, we also give an estimate on the dependence on β for the Sine_β process.

THEOREM 3. *Construct the* $\mathtt{Sine}_\beta$ *operators for all* $\beta > 0$ *with the same hyperbolic Brownian motion. Denote the eigenvalues corresponding to* β *by* $\{\lambda_{k,\beta}, k \in \mathbb{Z}\}$ *with* $\lambda_{0,\beta} < 0 \leq \lambda_{1,\beta}$. *Then for* $0 < \theta$, *there is an a.s. finite* $C = C_\theta$ *depending only on* θ *and* $\mathcal{B}$ *so that if* $\theta < \beta < \beta' \leq \infty$ *and* $\delta = \frac{4}{\beta} - \frac{4}{\beta'} \leq 1/3$ *then*

$$\sum_k \left(\frac{1}{\lambda_{k,\beta}} - \frac{1}{\lambda_{k,\beta'}}\right)^2 \leq \|\mathtt{rSine}_\beta - \mathtt{rSine}_{\beta'}\|_{\mathrm{HS}}^2 \leq C\delta \log(\delta^{-1}). \tag{5}$$

The theorem allows the choice $\beta' = \infty$. In this case, the driving path $\gamma(t)$ is just a constant, and the corresponding point process Sine_∞ is the so-called clock process: the set $2\pi\mathbb{Z}$ shifted by a uniformly distributed random variable on $[0, 2\pi]$. The bound (5) in this case provides a quantitative description of the limit in distribution of Sine_β as $\beta \to \infty$.

The theorem requires a positive lower bound on β, hence it cannot describe the $\beta \to 0$ behavior. It is known that in this case the limit in distribution of Sine_β is a homogeneous Poisson process; see Allez and Dumaz (2014).

Structure of the paper. The proof of Theorem 1 relies on a precise coupling of a hyperbolic random walk and hyperbolic Brownian motion.

The starting point is a hyperbolic heat kernel bound. Consider the squared Euclidean norm for a hyperbolic Brownian motion started at the origin in the Poincaré disk model. We show that this quantity at a small time is very close in total variation to a beta random variable, which also stochastically dominates it (Lemma 9 in Section 3.1).

This leads to a coupling of the hyperbolic random walk steps with hyperbolic Brownian increments at stopping times (Sections 3.2 and 3.3). For each single step, with high probability we stop at a fixed time. Otherwise, we just wait for the Brownian motion to hit the right distance. At the tail of the random walk, a slightly different coupling is implemented.

A modulus of continuity estimate for hyperbolic Brownian motion then implies that the random walk is close to the Brownian path (Section 3.4).

In Section 4, we show that if two paths are close and escape to the boundary of $\mathbb{H}$ similarly as a geodesic then the corresponding operators are also close. Finally, in Section 5 we use the linear rate of escape for hyperbolic Brownian motion to complete the proof of Theorem 1. Section 6 proves Theorem 3. Some of the technical facts needed are collected in the Appendix.

Historical background. The modern history of random matrices originates from Wigner (1951), who used them to approximate the spectrum of self-adjoint operators from statistical physics point of view.

In the following decades, the scaling behavior of a number of random matrix models were derived. The point process limits of the random matrix spectra were described via the limiting joint densities, usually relying on some algebraic structure of the finite models. (See the monographs Mehta (2004), Anderson, Guionnet and Zeitouni (2010) and Forrester (2010) for an overview of the classical results.)

Dumitriu and Edelman (2002) constructed tridiagonal random matrix models with spectrum distributed as beta ensembles, one parameter extensions of classical random matrix models. Edelman and Sutton (2007) observed that under the appropriate scaling, these tridiagonal matrix models behave like approximate versions of random stochastic operators, and conjectured that scaling limits of beta ensembles can be described as the spectra of these objects.

These conjectures were confirmed in Ramírez, Rider and Virág (2011) and Ramírez and Rider (2009) for the soft and hard edge scaling limits of beta ensembles. The authors rigorously defined the stochastic differential operators that show up as limits, and proved the convergence of the finite ensembles to the spectrum of these operators.

In Valkó and Virág (2009) and Killip and Stoiciu (2009), the bulk scaling limit of the Gaussian and circular beta ensembles were derived, and the counting functions of the limit processes were characterized via coupled systems of SDEs. In Nakano (2014) and Valkó and Virág (2017), it was shown that the scaling limit of the circular beta ensemble is the same as Sine_β, the bulk limit of the Gaussian beta ensemble. Furthermore, Valkó and Virág (2017) constructed a stochastic differential operator with a spectrum given by Sine_β and showed that several random matrix limits can be described via differential operators parametrized by certain random walks or diffusions.

The coupling of the circular beta ensemble for $\beta = 2$ (the circular unitary ensemble) to its limit, the Sine_2 process has been recently studied in Bourgade, Najnudel and Nikeghbali (2013), Maples, Najnudel and Nikeghbali (2019) and Meckes and Meckes (2016). In Bourgade, Najnudel and Nikeghbali (2013), the circular unitary ensembles of various sizes are coupled together and it is shown that the scaled ensembles converge a.s. to a Sine_2 process. Moreover, a bound of the form (4) is given with an exponent $\varepsilon > 0$. This coupling was further studied in Maples, Najnudel and Nikeghbali (2019) where a bound of the form (4) is given with an exponent $1/3$.

In Meckes and Meckes (2016), the total variation distance between the counting functions of the finite and the limiting process is considered. Denote by $\mathcal{N}_n$ the counting function of the appropriately scaled circular unitary ensemble of size n, and by $\mathcal{N}$ the counting function of the Sine_2 process. It is shown that for any fixed interval I the following bound holds:

$$d_{\mathrm{TV}}\big(\mathcal{N}_n(I), \mathcal{N}(I)\big) \le 5\frac{|I|^2}{n^{3/2}} \quad \text{for } n \ge n_0(I). \tag{6}$$

This provides a bound on the distance between the distributions for the number of points in a given interval, but does not seem to imply a process level result.

2. Stochastic differential operators.

We review the framework introduced in Valkó and Virág (2017) to study random matrix ensembles via differential operators.

2.1. *Dirac operators*.

We consider differential operators of the form

$$\tau : f \to R^{-1}(t) J \frac{d}{dt} f. \tag{7}$$

Here, $f : [0, 1) \to \mathbb{R}^2$, and

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad R = \frac{1}{2} X^t X, \qquad X = \frac{1}{\sqrt{y}} \begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix}, \tag{8}$$

with $x : [0, 1) \to \mathbb{R}$ and $y : [0, 1) \to (0, \infty)$. We consider boundary conditions parametrized by nonzero vectors $u_0, u_1 \in \mathbb{R}^2$ where we assume that $u_0^t J u_1 = 1$. We set the domain of the differential operator τ as

$$\mathrm{dom}(\tau) = \Big\{ v \in L_R^2 \cap \mathrm{AC} : \tau v \in L_R^2, v(0)^t J u_0 = 0, \lim_{s \to 1} v(s)^t J u_1 = 0 \Big\}. \tag{9}$$

Here, L_R^2 is the L^2 space of functions $f : [0, 1) \to \mathbb{R}^2$ with the L^2 norm $\|f\|_2^2 = \int_0^1 f^t R f \, ds$, while AC is the set of absolutely continuous functions.

The function $\gamma = x + iy$ is a path in the upper half-plane $\{(x, y) : y > 0\}$. In Valkó and Virág (2017), it was shown that various properties of τ can be identified by treating γ as a path in the hyperbolic plane $\mathbb{H}$ (using the upper half-plane representation) with u_0, u_1 identified with boundary points η_0, η_1 of $\mathbb{H}$. The set of boundary points of $\mathbb{H}$ in the upper half-plane representation is $\mathbb{R} \cup \{\infty\}$. A nonzero vector $v = (v_1, v_2)^t \in \mathbb{R}^2$ can be identified with the boundary point $\mathcal{P}v \in \partial\mathbb{H}$ where $\mathcal{P}v = \frac{v_1}{v_2}$ if $v_2 \neq 0$ and $\mathcal{P}v = \infty$ if $v_2 = 0$. To show the dependence on these parameters, we use the notation $\tau = \mathtt{Dir}(\gamma, \eta_0, \eta_1)$.

For a given boundary point $\eta \in \partial\mathbb{H}$, the (signed) horocyclic distance of points a and b in $\mathbb{H}$ with respect to η is defined as

$$d_\eta(a, b) = \lim_{z \to \eta} \big(d_{\mathbb{H}}(a, z) - d_{\mathbb{H}}(b, z) \big).$$

Here, $d_{\mathbb{H}}$ is the hyperbolic distance and the limit is evaluated along a sequence of points in $\mathbb{H}$ converging to η. We record the following formulas for the half-plane representation:

$$d_{\mathbb{H}}(x_1 + iy_1, x_2 + iy_2) = \operatorname{arccosh}\left(1 + \frac{(x_1 - x_2)^2 + (y_1 - y_2)^2}{2 y_1 y_2} \right), \tag{10}$$

$$d_\eta(x + iy, i) = \begin{cases} \log\left(\dfrac{1}{y}\right) & \text{if } \eta = \infty, \\ \log\left(\dfrac{(x - q)^2 + y^2}{(1 + q^2) y} \right) & \text{if } \eta = q \in \mathbb{R}. \end{cases} \tag{11}$$

The following theorem gives a condition in terms of the parameters γ, η_0, η_1 for τ to be self-adjoint with a Hilbert–Schmidt inverse.

THEOREM 4 (Valkó and Virág (2017)). *Let η_0, η_1 be distinct boundary points of $\mathbb{H}$ and $\gamma : [0, 1) \to \mathbb{H}$ be measurable and locally bounded. Assume that there is a $\xi \in \mathbb{H}$ with*

$$\int_0^1 e^{d_{\eta_1}(\gamma(t),\xi)}\,dt < \infty \quad \text{and} \quad \int_0^1 \int_s^1 e^{d_{\eta_0}(\gamma(s),\xi)+d_{\eta_1}(\gamma(s),\xi)}\,ds\,dt < \infty. \tag{12}$$

Then the operator $\tau = \mathtt{Dir}(\gamma, \eta_0, \eta_1)$ is self-adjoint on $\mathrm{dom}(\tau)$ *and its inverse is Hilbert–Schmidt. The inverse τ^{-1} is an integral operator on L^2_R with kernel function*

$$K(s,t) = \big(u_0 u_1^t \mathbf{1}(s<t) + u_1 u_0^t \mathbf{1}(s \ge t)\big). \tag{13}$$

This means that if $g \in L^2_R$ then $(\tau^{-1}g)(x) = \int_0^1 K(x,y)R(y)g(y)\,dy$.

Suppose that γ, η_0, η_1 satisfy the conditions of the theorem above and consider the operator $\tau = \mathtt{Dir}(\gamma, \eta_0, \eta_1)$. Let $\hat\tau = X\tau X^{-1}$; this means that $(\hat\tau f)(x) = X(x)(\tau g)(x)$ where $g(y) = X^{-1}(y)f(y)$. Then $\hat\tau$ is just τ after a change of coordinates. In particular, $\hat\tau$ is a self-adjoint differential operator on $\{v : X^{-1}v \in \mathrm{dom}(\tau)\} \subset L^2$, with the same spectrum as τ. We denote the inverse of $\hat\tau$ by $\mathtt{r}\tau$ ($\mathtt{r}$ standing for resolvent). By Theorem 4, the operator $\mathtt{r}\tau$ is an integral operator acting on L^2 functions with kernel

$$K_{\mathtt{r}\tau}(s,t) = \frac{1}{2}\big(a(s)c(t)^t \mathbf{1}(s<t) + c(s)a(t)^t \mathbf{1}(s \ge t)\big), \tag{14}$$

where $a(s) = X(s)u_0$ and $c(s) = X(s)u_1$. Thus for $g \in L^2$ we have

$$(\mathtt{r}\tau g)(x) = \int_0^1 K_{\mathtt{r}\tau}(x,y)g(y)\,dy.$$

2.2. *Stochastic operators*. The Gaussian and circular β ensembles are defined via the following joint densities on $\mathbb{R}^n$ and $[0, 2\pi)^n$, respectively,

$$p^{\mathrm{g}}_{\beta,n}(\lambda_1,\dots,\lambda_n) = \frac{1}{Z^{\mathrm{g}}_{n,\beta}} \prod_{1\le j<k\le n} |\lambda_j - \lambda_k|^\beta e^{-\frac{\beta}{4}\sum_{j=1}^n \lambda_j^2}, \tag{15}$$

$$p^{\mathrm{c}}_{\beta,n}(\lambda_1,\dots,\lambda_n) = \frac{1}{Z^{\mathrm{c}}_{n,\beta}} \prod_{1\le j<k\le n} \big|e^{i\lambda_j} - e^{i\lambda_k}\big|^\beta. \tag{16}$$

For $\beta = 2$, these give the joint eigenvalue densities of the Gaussian and circular unitary ensemble. We use angles to represent the eigenvalues in the circular case.

The bulk scaling limits of these ensembles have been identified in Valkó and Virág (2009) and Killip and Stoiciu (2009). In Nakano (2014) and Valkó and Virág (2017), it was shown that the scaling limit of the circular beta ensemble is the same as the bulk limit of the Gaussian beta ensemble.

THEOREM 5 (Valkó and Virág (2009)). *Fix $\beta > 0$ and $|E| < 2$. Let Λ^{g}_n be a finite point process with density* (15). *Then $\sqrt{4-E^2}\sqrt{n}(\Lambda^{\mathrm{g}}_n - \sqrt{n}E)$ converges in distribution to a point process* Sine_β.

THEOREM 6 (Killip and Stoiciu (2009), Nakano (2014), Valkó and Virág (2017)). *Fix $\beta > 0$ and let Λ^{c}_n be a finite point process with density* (16). *Then $n\Lambda^{\mathrm{c}}_n$ converges in distribution to the point process* Sine_β.

In Valkó and Virág (2017), the authors constructed random Dirac operators with spectrum given by Sine_β and the finite circular beta ensemble. Recall that the standard hyperbolic

Brownian motion $x + iy$ in the upper half-plane representation started from i is the solution of the SDE

$$d(x + iy) = y(dB_1 + i\,dB_2), \qquad x(0) + iy(0) = i,$$

where B_1, B_2 are independent standard real Brownian motions.

THEOREM 7 (Valkó and Virág (2017)). *Fix $\beta > 0$ and let $\mathcal{B}$ be a standard hyperbolic Brownian motion in the upper half-plane started from $\mathcal{B}(0) = i$. Set $\tilde{\mathcal{B}}(t) = \mathcal{B}(-\frac{4}{\beta}\log(1 - t))$, $t \in [0, 1)$, $\eta_0 = \infty$ and $\eta_1 = \lim_{t\to\infty} \mathcal{B}(t)$. Then the operator*

$$\mathtt{Sine}_\beta = \mathtt{Dir}\big(\tilde{\mathcal{B}}(t), \eta_0, \eta_1\big) \tag{17}$$

is a.s. self-adjoint with a Hilbert–Schmidt inverse and $\operatorname{spec}(\mathtt{Sine}_\beta) \overset{d}{=} \mathrm{Sine}_\beta$.

The definition (17) can be extended to $\beta = \infty$. In this case, we set $\tilde{\mathcal{B}}(t) = i$ for $t \in [0, 1)$. Then the corresponding operator is just $2J\frac{d}{dt}$ on $[0, 1)$ with boundary conditions $u_0 = (-1, 0)^t$ and $u_1 = (1, \eta_1)^t$, with the same η_1 as in Theorem 7. We call this operator $\mathtt{Sine}_\infty$. A simple computation shows that $\operatorname{spec}(\mathtt{Sine}_\infty) \overset{d}{=} 2\pi\mathbb{Z} + U$ where U is uniform on $[0, 2\pi]$. We denote this process Sine_∞, this is sometimes referred to as the clock process.

Killip and Nenciu (2004) gave a construction for generating a random unitary matrix with eigenvalues distributed as the circular beta ensemble. Building on this result, Valkó and Virág (2017) produced a random Dirac operator representation for the circular beta ensemble using a random walk in $\mathbb{H}$.

Fix $n \ge 1$ and let $\zeta_0, \dots, \zeta_{n-2}$ be independent with ζ_k distributed as $\mathrm{Beta}(1, \frac{\beta}{2}(n-k-1))$.[1] Set $Y_k = \log(\frac{1+\sqrt{\zeta_k}}{1-\sqrt{\zeta_k}})$. We define the random walk $b_0, b_1, \dots, b_{n-1}$ in $\mathbb{H}$ with a final boundary point $b_n \in \partial\mathbb{H}$. We set $b_0 = i$, and for $0 \le k \le n-2$ we choose b_{k+1} uniformly (according to the hyperbolic geometry) among the points in $\mathbb{H}$ with hyperbolic distance Y_k from b_k, independently of the previous choices. The final point b_n is chosen uniformly on the boundary $\partial\mathbb{H}$ as viewed from b_{n-1}, independently of the previous choices. Note that ζ_k is the squared Euclidean norm of the random walk step in the Poincaré disk model with b_k at the origin.

THEOREM 8 (Valkó and Virág (2017)). *Consider the random walk $b_0, b_1, \dots, b_n$ defined above. Set $\eta_0 = \infty$, $\eta_1 = b_n$ and $\mathcal{B}_n(t) = b_{\lfloor nt\rfloor}$ for $t \in [0, 1)$. The operator*

$$\mathtt{Circ}_{\beta,n} = \mathtt{Dir}\big(\mathcal{B}_n(t), \eta_0, \eta_1\big)$$

is a.s. self-adjoint with a Hilbert–Schmidt inverse and $\operatorname{spec}(\mathtt{Circ}_{\beta,n}) \overset{d}{=} n\Lambda_n^{\mathrm{c}} + 2\pi n\mathbb{Z}$ *where Λ_n^{c} is the finite point process with joint density* (16).

Theorems 7 and 8 imply that zero is not an eigenvalue for the operators $\mathtt{Sine}_\beta$ and $\mathtt{Circ}_{\beta,n}$ with probability one.

2.3. *Coupling $\mathtt{Circ}_{\beta,n}$ and $\mathtt{Sine}_\beta$.* Fix $\beta > 0$. Let $\mathcal{B}$ be a hyperbolic Brownian motion, $\eta_0 = \infty$, $\eta_1 = \mathcal{B}(\infty)$, and consider the operator $\mathtt{Sine}_\beta = \mathtt{Dir}(\mathcal{B}(-\frac{4}{\beta}\log(1-t)), \eta_0, \eta_1)$.

Let U be a uniform random variable on $[0, 1]$ independent from $\mathcal{B}$, and let $\mathcal{F}_t, t \ge 0$ be the natural filtration of $\mathcal{B}$ enlarged with U. In Proposition 12 below, we construct an

[1] A random variable has distribution $\mathrm{Beta}(a, b)$ with $a, b > 0$ if it has density $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}\mathbf{1}(x \in [0, 1])$.

array of stopping times $\tau_{n,k}$ with respect to $\mathcal{F}_t$ so that for each $n \geq 1$ the random variables $\mathcal{B}(\tau_{n,n}), \mathcal{B}(\tau_{n,n-1}), \dots, \mathcal{B}(\tau_{n,0})$ have the same joint distribution as the random walk $b_0, \dots, b_n$ from Theorem 8. (Note that $\mathcal{B}(\tau_{n,k})$ corresponds to b_{n-k}, so the index k matches up with the parameter of the Beta distribution in the appropriate step of the random walk.) Setting $\mathtt{Circ}_{\beta,n} = \mathtt{Dir}(\tilde{\mathcal{B}}_n(t), \eta_0, \eta_1)$ with $\tilde{\mathcal{B}}_n(t) = \mathcal{B}(\tau_{n,\lceil (1-t)n \rceil}), t \in [0, 1)$ gives the appropriate coupling of $\mathtt{Sine}_\beta$ and the sequence $\{\mathtt{Circ}_{\beta,n}\}_{n\geq 1}$ that is used in Theorem 1.

3. Coupling construction. The goal of this section is to construct the coupling of the hyperbolic random walk and the hyperbolic Brownian motion that appears in Theorem 1. Since the steps in the random walk have rotationally invariant distributions and the same is true for the increments of the hyperbolic Brownian motion, it would be easy to embed the walk via simple hitting times. However, this "naive" embedding would not give enough control for us to obtain the error bound in Theorem 1. Instead we construct a coupling that also exploits the fact that the single step hyperbolic distance distributions in the random walk can be well approximated with the distance distribution of the hyperbolic Brownian motion at a certain fixed time.

3.1. *A heat kernel bound on the hyperbolic plane.* Our coupling relies on a careful estimate of the transition density of hyperbolic Brownian motion. Although there are a number of similar bounds in the literature (see, e.g., Davies and Mandouvalos (1988)), we could not find one that would be strong enough for our purposes. We show that the distribution of the distance of hyperbolic Brownian motion from its starting point at time $t \leq 1$ can be well approximated estimated using a Beta$(1, \frac{2}{t} - 1/2)$ random variable.

LEMMA 9. *Let $\mathcal{B}(t)$ be standard hyperbolic Brownian motion and let $t \in (0, 1]$. Let $Y = \log(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}})$ where ξ has distribution* Beta$(1, \frac{2}{t} - 1/2)$ *and set $\zeta = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$. Then the following statements hold*:

(a) *$P(Y > r) \geq P(\zeta > r)$ for all $r > 0$, in other words Y stochastically dominates ζ.*
(b) *The total variation distance of ζ and Y is bounded by $\frac{3}{2}t$.*

The proof of the lemma relies on a precise analysis of the explicit formula for the transition density. We leave it for Section A.1 in the Appendix.

3.2. *Single step coupling.* We first concentrate on a single step in the hyperbolic random walk corresponding to $\mathtt{Circ}_{\beta,n}$ and couple it to the hyperbolic Brownian motion.

PROPOSITION 10. *Fix $\gamma \geq 3/2$. Let $\mathcal{B}$ be hyperbolic Brownian motion and U an independent uniform random variable on* $[0, 1]$. *Let $\mathcal{F}_t, t \geq 0$ be the filtration of $\mathcal{B}$ enlarged with U. Consider a Poincaré disk representation of the hyperbolic plane where $\mathcal{B}(0) = 0$.*

There exists a finite random variable $\sigma > 0$ so that the following hold:

1. *σ is a stopping time with respect to $\mathcal{F}_t$.*
2. *$\mathcal{B}(\sigma)$ has rotationally invariant distribution and $|\mathcal{B}(\sigma)|^2$ has* Beta$(1, \gamma)$ *distribution.*
3. *$P(\sigma \geq \frac{4}{2\gamma+1}) = 1$ and $P(\sigma \neq \frac{4}{2\gamma+1}) \leq \frac{3}{\gamma}$.*
4. *For $r \geq 8$, we have $P(\sigma > r/\gamma) \leq 3e^{-\frac{1}{5}r^{1/3}}$.*

The proof of the proposition will rely on Lemma 9 and the Lemma 11 below. Lemma 11 is a standard coupling statement, we include its proof in the Appendix for completeness.

LEMMA 11. *Assume that X_1 and X_2 are random variables so that X_2 stochastically dominates X_1 and the total variation distance of their distributions is ε. Then there exists a measurable function $g : \mathbb{R}^2 \to \mathbb{R}$ so that if U is a uniform random variable on* $[0, 1]$, *independent of X_1 then the following hold*:

(a) *$g(X_1, U)$ has the same distribution as X_2.*
(b) *$P(X_1 \le g(X_1, U)) = 1$.*
(c) *$P(X_1 = g(X_1, U)) = 1 - \varepsilon$.*

PROOF OF PROPOSITION 10. Set $t = \frac{4}{2\gamma+1}$, then $0 < t < 1$. Recall that if z is in the Poincaré disk with $|z| = r < 1$ then $d_{\mathbb{H}}(0, z) = \log(\frac{1+r}{1-r})$.

Let ξ be a random variable with distribution $\mathrm{Beta}(1, \gamma)$. Then by Lemma 9 the random variable $\log(\frac{1+|\mathcal{B}(t)|}{1-|\mathcal{B}(t)|})$ is stochastically dominated by $\log(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}})$ and their total variation distance is bounded by $\frac{3}{2}t$. Since $\log(\frac{1+r}{1-r})$ is strictly increasing in r, we get that $|\mathcal{B}(t)|$ is stochastically dominated by $\sqrt{\xi}$ and their total variation distance is bounded by $\frac{3}{2}t$.

By Lemma 11, there exits a measurable function g so that almost surely $g(|\mathcal{B}(t)|, U) \ge |\mathcal{B}(t)|$, $g(|\mathcal{B}(t)|, U)$ has the same distribution as $\sqrt{\xi}$, and $P(|\mathcal{B}(t)| \ne g(|\mathcal{B}(t)|, U)) \le \frac{3}{2}t \le 3/\gamma$.

We set

$$\sigma = \inf\{s \ge t : |\mathcal{B}(s)| = g(|\mathcal{B}(t)|, U)\}.$$

Then σ is an a.s. finite stopping time with respect to $\mathcal{F}_t$ and almost surely $\sigma \ge t$. Because σ only depends on $|\mathcal{B}|$ and U, it follows that $\mathcal{B}(\sigma)$ has rotationally invariant distribution. Finally, from our construction we get that $|\mathcal{B}(\sigma)|^2 = g(|\mathcal{B}(t)|, U)^2$ has $\mathrm{Beta}(1, \gamma)$ distribution and $P(\sigma \ne t) \le \frac{3}{2}t \le 3/\gamma$.

The only thing left to prove is the tail bound for σ. We start with the bound

$$\begin{aligned} P(\sigma > r/\gamma) \le {} & P\biggl(\sigma > r/\gamma \text{ and } d_{\mathbb{H}}(0, \mathcal{B}(\sigma)) \le \frac{r^{1/3}}{\gamma^{1/2}}\biggr) \\ & + P\biggl(d_{\mathbb{H}}(0, \mathcal{B}(\sigma)) > \frac{r^{1/3}}{\gamma^{1/2}}\biggr). \end{aligned} \tag{18}$$

For the rest of the proof, we assume $r \ge 8$. Then $\sigma > r/\gamma > t$ and from the definition of σ, it follows that $d_{\mathbb{H}}(0, \mathcal{B}(s)) < d_{\mathbb{H}}(0, \mathcal{B}(\sigma))$ for $t \le s < \sigma$. Thus we can bound the first term on the right of (18) by writing

$$\begin{aligned} & P\biggl(\sigma > r/\gamma \text{ and } d_{\mathbb{H}}(0, \mathcal{B}(\sigma)) \le \frac{r^{1/3}}{\gamma^{1/2}}\biggr) \\ & \quad \le P\biggl(\max_{t \le s \le r/\gamma} d_{\mathbb{H}}(0, \mathcal{B}(s)) \le \frac{r^{1/3}}{\gamma^{1/2}}\biggr) \\ & \quad \le P\biggl(\max_{t \le s \le r/\gamma} d_{\mathbb{H}}(\mathcal{B}(t), \mathcal{B}(s)) \le 2\frac{r^{1/3}}{\gamma^{1/2}}\biggr) \\ & \quad = P\biggl(\max_{0 \le s \le r/\gamma - t} d_{\mathbb{H}}(0, \mathcal{B}(s)) \le 2\frac{r^{1/3}}{\gamma^{1/2}}\biggr). \end{aligned}$$

Since $t \le \frac{3}{2\gamma}$ and $r \ge 8$, we have $r/\gamma - t \ge \frac{4}{5} r\gamma^{-1}$. Using the bound (67) from Lemma 20 of the Appendix, we get

$$P\left(\max_{0\le s\le r/\gamma-t} d_{\mathbb{H}}(0,\mathcal{B}(s)) \le 2\frac{r^{1/3}}{\gamma^{1/2}}\right) \le P\left(\max_{0\le s\le \frac{4}{5}r\gamma^{-1}} d_{\mathbb{H}}(0,\mathcal{B}(s)) \le 2\frac{r^{1/3}}{\gamma^{1/2}}\right)$$
$$\le \frac{4}{\pi} e^{-\frac{\pi^2 r^{1/3}}{40}}.$$

For the second term in (18), we recall that by construction $d_{\mathbb{H}}(0,\mathcal{B}(\sigma))$ has the same distribution as $\log(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}})$ where ξ has distribution $\mathrm{Beta}(1,\gamma)$. By an explicit computation,

$$P\left(\log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right) > u\right) = \operatorname{sech}^{2\gamma}\left(\frac{u}{2}\right).$$

We have $\operatorname{sech}(x) = \frac{2}{e^x+e^{-x}} \le 2e^{-x}$ for $x > 0$, which gives the following upper bound for $u \ge 4\log 2$:

$$P\left(\log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right) > u\right) \le 2^{2\gamma} e^{-\gamma u} \le e^{-\frac{\gamma u}{2}}.$$

For $0 \le u \le 4\log 2$, we have $\log\operatorname{sech}(u/2) \le -\frac{u^2}{12}$ so for these values we get

$$P\left(\log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right) > u\right) \le e^{-\frac{\gamma u^2}{6}}. \tag{19}$$

From this, we get

$$P\left(d_{\mathbb{H}}(0,\mathcal{B}(\sigma)) > \frac{r^{1/3}}{\gamma^{1/2}}\right) \le \max\left(e^{-\frac{\sqrt{\gamma}r^{1/3}}{2}}, e^{-\frac{r^{2/3}}{6}}\right) \le e^{-\frac{1}{3}r^{1/3}},$$

where in the last step we used $\gamma \ge 3/2$ and $r \ge 8$. Collecting our estimates, we get

$$P(\sigma > r/\gamma) \le \frac{4}{\pi} e^{-\frac{\pi^2 r^{1/3}}{40}} + e^{-\frac{1}{3}r^{1/3}} \le 3e^{-\frac{1}{5}r^{1/3}},$$

which completes the proof of the proposition. □

3.3. *Path coupling*. Using Proposition 10 repeatedly, we can provide a coupling of the hyperbolic Brownian motion and the hyperbolic random walk appearing in the construction of $\mathtt{Circ}_{\beta,n}$.

PROPOSITION 12. *Let $\mathcal{B}$ be hyperbolic Brownian motion, $U_k, k \ge 1$* Uniform[0, 1] *random variables, and $\xi_k, k \ge 1$ random variables with distribution* $\mathrm{Beta}(1, \frac{\beta}{2}k)$, *with $\mathcal{B}, U_1, U_2, \ldots, \xi_1, \xi_2, \ldots$ all independent. Let $\mathcal{F}_t, t \ge 0$ be the filtration of $\mathcal{B}$ enlarged with the random variables $U_k, \xi_k, k \ge 1$.*

There exists a collection of stopping times $(\tau_{n,k}; 1 \le n, 0 \le k \le n)$ with respect to $\mathcal{F}_t, t \ge 0$ so that the following statements hold:

1. *For each fixed n, we have $0 = \tau_{n,n} < \tau_{n,n-1} < \cdots < \tau_{n,0} = \infty$, and the random variables $\Delta\tau_{n,k} := \tau_{n,k} - \tau_{n,k+1}$ are independent for $k = 1, 2, \ldots, n-1$.*

2. *For each n, the process $(b_k^{(n)} = \mathcal{B}(\tau_{n,n-k}), k = 0, 1, \dots, n)$ is a hyperbolic random walk with the same distribution as the one given above Theorem 8.*

3. *Let $t_{n,k} = \frac{4}{\beta}\log(\frac{n}{k})$ for $1 \le k \le n$. There exists a random integer $N_0 > 0$ so that for $n \ge N_0$ and $\log^6 n \le k \le n$ we have almost surely*

$$t_{n,k} - \frac{4}{\beta^2 k} \le \tau_{n,k} \le t_{n,k} + \frac{\log^{4+1/2} n}{k}. \tag{20}$$

4. *For k fixed the hyperbolic distance $d_{\mathbb{H}}(\mathcal{B}(\tau_{n,k}), \mathcal{B}(\tau_{n,k-1}))$ does not depend on n as long as $k < \log^6 n$.*

PROOF. We first give the construction of the stopping times, then prove that they satisfy all the conditions.

For a fixed $n \ge 1$, we define $\tau_{n,k}$ recursively, starting with $\tau_{n,n} = 0$. If for a certain $k < n$, we have already defined $\tau_{n,k+1}$ then we define $\tau_{n,k}$ as follows.

- *In the case of $k \ge \max(\log^6 n, \frac{3}{\beta})$:*

  We apply Proposition 10 with $\gamma = \frac{\beta}{2}k$ for the hyperbolic Brownian motion $\tilde{\mathcal{B}}(t) = \mathcal{B}(t + \tau_{n,k+1}) - \mathcal{B}(\tau_{n,k+1}), t \ge 0$ and the independent uniform random variable U_k, and denote the constructed stopping time by $\sigma_{n,k}$. We set $\tau_{n,k} = \tau_{n,k+1} + \sigma_{n,k}$.
- *In the case of $1 \le k < \max(\log^6 n, \frac{3}{\beta})$:*

  We set

$$\tau_{n,k} = \inf\left\{t \ge \tau_{n,k+1} : d_{\mathbb{H}}\big(\mathcal{B}(\tau_{n,k+1}), \mathcal{B}(t)\big) = \log\left(\frac{1+\sqrt{\xi_k}}{1-\sqrt{\xi_k}}\right)\right\}.$$

- For $k = 0$, we define $\tau_{n,k} = \infty$.

Note that we use the coupling given in Proposition 10 when k is not too small compared to n. This will enable us to prove the estimate (20), and allows us to control the distance between $\mathcal{B}(t_{n,k})$ and $b_k^{(n)}$ (see Proposition 13 below) in this regime. The bounds in Proposition 10 are not strong enough to use this approach for all k, that is why we need to use a different coupling construction for small values of k.

By construction, the random variables $\tau_{n,k}$, $1 \le k \le n$ are a.s. finite stopping times with respect to the filtration $\mathcal{F}_t, t \ge 0$, and they satisfy conditions 1, 2 and 4. To check Condition 3, we first choose n_0 so that for $n \ge n_0$ we have $\log^6 n \ge \frac{3}{\beta}$ and $\log^{3+3/8} n \ge 8$. During the rest of this proof, we will assume $1 \le \log^6 n \le k \le n$.

For $n \ge n_0$, by Proposition 10 we have almost surely

$$\tau_{n,k} \ge \sum_{j=k}^{n-1} \frac{4}{\beta j + 1} \ge \frac{4}{\beta}\log\left(\frac{n}{k}\right) - \frac{4}{\beta^2}\frac{1}{k}.$$

This takes care of the lower bound in (20).

For the upper bound, recall the definition of $\sigma_{n,k} = \tau_{n,k} - \tau_{n,k+1}$. From Proposition 10, we have the following estimates:

$$\begin{gathered} P\left(\sigma_{n,k} \ge \frac{4}{\beta k + 1}\right) = 1, \qquad P\left(\sigma_{n,k} = \frac{4}{\beta k + 1}\right) \ge 1 - \frac{6}{\beta k}, \\ P\left(\sigma_{n,k} > \frac{\log^{3+3/8} n}{\frac{\beta}{2}k}\right) \le 3e^{-\frac{1}{5}\log^{1+1/8} n}. \end{gathered} \tag{21}$$

Since $\sum_{k,n} P(\sigma_{n,k} > \frac{\log^{3+3/8} n}{\frac{\beta}{2}k}) \le 3\sum_n ne^{-\frac{1}{5}\log^{1+1/8} n} < \infty$, by the Borel–Cantelli lemma there is a random $N_1 \ge n_0$ so that for $n \ge N_1$ we have $\sigma_{n,k} \le \frac{\log^{3+3/8} n}{\frac{\beta}{2}k}$ a.s. Set

$$Z_{n,k} = \frac{4}{\beta k+1}\mathbf{1}\left(\sigma_{n,k} = \frac{4}{\beta k+1}\right) + \frac{2\log^{3+3/8} n}{\beta k}\mathbf{1}\left(\sigma_{n,k} \ne \frac{4}{\beta k+1}\right). \tag{22}$$

For $n \ge N_1$, we have $\sigma_{n,k} \le Z_{n,k}$ and $\tau_{n,k} \le \sum_{j=k}^{n-1} Z_{n,j}$ a.s.

Next, we will bound $P(\sum_{j=k}^{n-1}(Z_{n,j} - \frac{4}{\beta j+1}) \ge \frac{\log^{9/2} n}{k})$. For $\lambda > 0$ from (21) and (22), we get

$$E\left(e^{\sum_{j=k}^{n-1}\lambda(Z_{n,j}-\frac{4}{\beta k+1})}\right) \le \prod_{j=k}^{n-1}\left(1 - \frac{6}{\beta j} + \frac{6}{\beta j}e^{\lambda\frac{2\log^{3+3/8} n}{\beta j}}\right). \tag{23}$$

Assuming that $\lambda\frac{2\log^{3+3/8} n}{\beta k} \le 1$, we can use that $\frac{e^x-1}{x} \le 2$ for $x \le 1$ to bound the right-hand side of (23) as

$$\begin{aligned}
\prod_{j=k}^{n-1}\left(1 - \frac{6}{\beta j} + \frac{6}{\beta j}e^{\lambda\frac{2\log^{3+3/8} n}{\beta j}}\right) &\le \prod_{j=k}^{n-1}\left(1 + \frac{6}{\beta j}\cdot 2\lambda\frac{2\log^{3+3/8} n}{\beta j}\right)\\
&\le \prod_{j=k}^{n-1} e^{\frac{24\lambda\log^{3+3/8} n}{\beta^2 j^2}}\\
&\le e^{\frac{48\lambda\log^{3+3/8} n}{\beta^2 k}}.
\end{aligned}$$

Setting now $\lambda = \frac{\beta k}{2\log^{3+3/8} n}$ and using the exponential Markov inequality, we obtain

$$\begin{aligned}
&P\left(\sum_{j=k}^{n-1}\left(Z_{n,j} - \frac{4}{\beta j+1}\right) \ge \frac{\log^{4+1/2} n}{k}\right)\\
&\quad = E\left(e^{\sum_{j=k}^{n-1}\lambda(Z_{n,j}-\frac{4}{\beta k+1})}\right)e^{-\lambda\frac{\log^{4+1/2} n}{k}}\\
&\quad \le e^{\frac{48\lambda\log^{3+3/8} n}{\beta^2 k} - \lambda\frac{\log^{4+1/2} n}{k}} \le e^{\frac{24}{\beta} - \frac{\beta}{2}\log^{1+1/8} n}.
\end{aligned}$$

Since $\sum_n ne^{\frac{24}{\beta}-\frac{\beta}{2}\log^{1+1/8} n} < \infty$, the Borel–Cantelli lemma implies that there is a random $N_0 \ge N_1$ so that for $n \ge N_0$ and $\log^6 n \le k \le n$ we have a.s.,

$$\sum_{j=k}^{n-1} Z_{n,j} \le \sum_{j=k}^{n-1}\frac{4}{\beta j+1} + \frac{\log^{4+1/2} n}{k} \le \frac{4}{\beta}\log\left(\frac{n}{k}\right) + \frac{\log^{4+1/2} n}{k}.$$

Since $\tau_{n,k} \le \sum_{j=k}^{n-1} Z_{n,j}$, the upper bound in (20) follows. □

3.4. *Path comparison.* Let $\mathcal{B}$ be a hyperbolic Brownian motion and consider the stopping times $\tau_{n,k}$ constructed in Proposition 12. Set $\tilde{\mathcal{B}}_n(t) = \mathcal{B}(\tau_{n,\lceil(1-t)n\rceil})$ and $\tilde{\mathcal{B}}(t) = \mathcal{B}(-\frac{4}{\beta}\log(1-t))$. The next proposition gives uniform bounds on $d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t))$. The estimates rely on path properties of the hyperbolic Brownian motion which are stated in Propositions 22 and 23, and proved in the Appendix.

PROPOSITION 13. *There is a random integer N^* so that for $n \geq N^*$ we have the following a.s. inequalities*:

$$d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t)) \leq \frac{\log^{3-1/8} n}{\sqrt{(1-t)n}} \quad \textit{if } 0 \leq t \leq 1 - \frac{\log^6 n}{n}, \tag{24}$$

$$d_{\mathbb{H}}\left(\tilde{\mathcal{B}}_n\left(1 - \frac{\log^6 n}{n}\right), \tilde{\mathcal{B}}_n(t)\right) \leq (\log\log n)^4 \quad \textit{if } 1 - \frac{\log^6 n}{n} \leq t < 1. \tag{25}$$

PROOF. Let $T_n = 1 - \frac{\log^6 n}{n}$. Consider N_0 from the statement of Proposition 12. For $n \geq N_0$ and $0 \leq t \leq T_n$, we have a.s.,

$$\begin{aligned} &\left|\tau_{n,\lceil (1-t)n\rceil} - \frac{4}{\beta}\log\left(\frac{1}{1-t}\right)\right| \\ &\quad \leq \frac{4}{\beta^2\lceil (1-t)n\rceil} + \frac{\log^{9/2} n}{\lceil (1-t)n\rceil} + \frac{4}{\beta}\left|\log\left(\frac{n}{\lceil (1-t)n\rceil}\right) - \log\left(\frac{1}{1-t}\right)\right| \\ &\quad \leq \frac{\log^{4+5/8} n}{(1-t)n}, \end{aligned} \tag{26}$$

where for the second inequality we also assume $n \geq n_0$ with an n_0 only depending on β. Inequality (26) implies

$$d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t)) \leq \max\{d_{\mathbb{H}}(\mathcal{B}(s), \mathcal{B}(s+u)) : |u| \leq h, 0 \leq s+u\},$$

with $h = \frac{\log^{4+5/8} n}{(1-t)n}$ and $s = \frac{4}{\beta}\log(\frac{1}{1-t})$. Note that $h \leq \frac{\log^{4+5/8} n}{\log^6 n} \leq \frac{1}{\log n}$.

Consider the random constant h_0 from the statement of Proposition 22 of the Appendix. If $n \geq e^{h_0^{-1}}$, then $h \leq h_0$ and we may apply Proposition 22 with $s, s+u$ if $0 \leq u \leq h$ and with $s+u, s$ if $0 \leq -u \leq \min(h, s)$. Using the fact that $h\log(2 + \frac{s+1}{h})$ is monotone increasing in h and s, we get

$$\begin{aligned} &\max\{d_{\mathbb{H}}(\mathcal{B}(s), \mathcal{B}(s+u)) : |u| \leq h, 0 \leq s+u\} \\ &\quad \leq 20\sqrt{h\log\left(2 + \frac{s+1}{h}\right)} \leq \frac{\log^{3-1/8} n}{\sqrt{(1-t)n}}, \end{aligned}$$

if $0 \leq t \leq T_n$ and $n \geq N_1$, with a random integer N_1.

Next, we prove the estimate for the $T_n \leq t < 1$ case. Recall the construction of the stopping times $\tau_{n,k}$ from the proof of Proposition 12. From the construction, it follows that

$$\begin{aligned} &d_{\mathbb{H}}\left(\tilde{\mathcal{B}}_n\left(1 - \frac{\log^6 n}{n}\right), \tilde{\mathcal{B}}_n(t)\right) \\ &\quad \leq \max\{d_{\mathbb{H}}(\mathcal{B}(\tau_{n,\lceil (1-T_n)n\rceil}), \mathcal{B}(s)) : \tau_{n,\lceil (1-T_n)n\rceil} \leq s \leq \tau_{n,1}\}. \end{aligned} \tag{27}$$

From Proposition 22 (and the comment after it), we get that there is a random constant $C_{\mathcal{B}}$ depending only on $\mathcal{B}$ so that

$$\begin{aligned} &\max\{d_{\mathbb{H}}(\mathcal{B}(\tau_{n,\lceil (1-T_n)n\rceil}), \mathcal{B}(s)) : \tau_{n,\lceil (1-T_n)n\rceil} \leq s \leq \tau_{n,1}\} \\ &\quad \leq C_{\mathcal{B}}(\tau_{n,1} - \tau_{n,\lceil (1-T_n)n\rceil})\sqrt{\log(3 + \tau_{n,1})}. \end{aligned} \tag{28}$$

We will show that there is a random constant $0 < C < \infty$ (not depending on n) so that a.s. for all n

$$\tau_{n,1} - \tau_{n,\lceil (1-T_n)n\rceil} \leq C(\log\log n)^3, \quad \tau_{n,1} \leq C\log n. \tag{29}$$

This bound together with (27) and (28) implies the estimate (25) for $n > N^*$, where N^* is random.

Consider the Beta distributed random variables ξ_j used in the construction in the proof of Proposition 12. Setting $Y_j = \log(\frac{1+\sqrt{\xi_j}}{1-\sqrt{\xi_j}})$ and using the tail bound (19), we get the following bound for $0 < r < 4\log 2$:

$$P(Y_j \ge r) = \operatorname{sech}^{\beta j}(r/2) \le e^{-\frac{\beta j r^2}{12}}.$$

This implies $\sum_j P(Y_j \ge \frac{4}{\sqrt{\beta}}\sqrt{\frac{\log 2j}{j}}) < \infty$, and shows that there is a random constant $C_\xi < \infty$ depending only on the sequence $\{\xi_n\}$, so that a.s.,

$$Y_j < C_\xi \sqrt{\frac{\log 2j}{j}}. \tag{30}$$

We will prove that there are random constants A and N_1 so that for all $n > N_1$ and $1 \le k < n(1 - T_n)$ we have a.s.,

$$\tau_{n,k} - \tau_{n,k+1} \le A \frac{\log(2k)\log\log n}{k}. \tag{31}$$

The bound (26) applied for $t = T_n$ shows that for $n \ge N_0$,

$$\tau_{n,\lceil(1-T_n)n\rceil} \le \frac{4}{\beta}\log n. \tag{32}$$

From (31) and (32), the bounds in (29) follow directly.

Let C_0 be the random constant from the statement of Proposition 23. Choose $A > 0$ so that $\frac{C_0 A}{2} > C_\xi^2$ where C_ξ is the constant from (30), and set $h_{n,k} = A\frac{\log(2k)\log\log n}{k}$. For $1 \le k < n(1 - T_n)$, we have

$$A\frac{\log\log n}{\log^6 n} \le h_{k,n} \le A\frac{\log(2\log^6 n)\log\log n}{k}$$

for $n \ge 3$. From this, it follows that there is a constant $N_2 \ge 3$ (depending only on C_0 and C_ξ) so that the following inequality holds for all $N_2 \le n$ and $1 \le k < n(1 - T_n) = \log^6 n$:

$$\frac{C_0}{4}\frac{h_{n,k}}{\log(2 + \frac{8}{\beta}\log n + h_{n,k}) + \log(h_{n,k} + h_{n,k}^{-1})} \ge C_\xi^2 \frac{\log 2k}{k}. \tag{33}$$

Consider $1 \le k < n(1 - T_n)$ and the stopping time $\tau_{n,k+1}$. By (74) of Proposition 23, there exists $0 < u < h_{k,n}$ so that

$$\begin{aligned} &d_{\mathbb{H}}(\mathcal{B}(\tau_{n,k+1}), \mathcal{B}(\tau_{n,k+1} + u))^2 \\ &\quad \ge \frac{C_0}{4}\frac{h_{n,k}}{\log(2 + \tau_{n,k+1} + h_{n,k}) + \log(h_{n,k} + h_{n,k}^{-1})}. \end{aligned}$$

If

$$\frac{C}{4}\frac{h_{n,k}}{\log(2 + \tau_{n,k+1} + h_{n,k}) + \log(h_{n,k} + h_{n,k}^{-1})} \ge C_\xi^2 \frac{\log 2k}{k} \tag{34}$$

then by (30) and the construction of the stopping times $\tau_{n,k}$ we had $\tau_{n,k} \le \tau_{n,k+1} + h_{n,k}$. The inequality (34) holds for $k + 1 = \lceil n(1 - T_n)\rceil$ if $n \ge N_2$ by (32) and (33). From this (34), and hence $\tau_{n,k} \le \tau_{n,k+1} + h_{n,k}$ follows for all $1 \le k < n(1 - T_n)$ by induction as long as $\sum_{k=1}^{n(1-T_n)} h_{k,n} \le \frac{4}{\beta}\log n$, which holds for n large enough. This completes the proof of (31) and also that of (25). $\square$

4. Hilbert–Schmidt bounds. This section contains general bounds on Dirac operators in the case when the corresponding hyperbolic paths escape to the boundary of $\mathbb{H}$ with a positive speed. Proposition 14 below shows that such a Dirac operators has a Hilbert–Schmidt inverse. The next result, Proposition 15 below, compares such a Dirac operator to its truncated version. Finally, Proposition 16 below compares two Dirac operators if their driving paths are close.

These propositions will be used in the next section to prove Theorems 1 and 3. In particular, for Theorem 1 we will bound $\|\mathtt{rSine}_\beta - \mathtt{rCirc}_{\beta,n}\|_{\mathrm{HS}}^2$ by replacing each integral operator with its truncated version using Proposition 15, and then use the path comparison in Proposition 13 together with Proposition 16 to estimate the norm difference of the truncated operators.

PROPOSITION 14 (Hilbert–Schmidt property). *Let $\{\gamma(t), 0 \le t\}$ be a measurable path in $\mathbb{H}$, $\eta_0, \eta_1 \in \partial\mathbb{H}$ distinct boundary points and $z_0 \in \mathbb{H}$. For a $\nu > 0$, set $\tilde{\gamma}(t) = \gamma(\nu \log(\frac{1}{1-t}))$. Let $z(t)$ be the point moving with speed $\alpha > 0$ on the geodesic connecting z_0 to η_1 with $z(0) = z_0$. Assume that there are constants $b > 0$ and $0 \le \varepsilon < \nu^{-1}$ so that for all $t \ge 0$ we have*

$$d_{\mathbb{H}}\big(\gamma(t), z(t)\big) \le b + \varepsilon t. \tag{35}$$

Then the operator

$$\tau = \mathtt{Dir}\big(\tilde{\gamma}(t), \eta_0, \eta_1\big) \tag{36}$$

is self-adjoint on the appropriate domain and τ^{-1} is Hilbert–Schmidt.

Here, the "appropriate domain" is described in and around (9).

PROOF. We will check that the conditions of Theorem 4 are satisfied.

If Q is an isometry of $\mathbb{H}$, then $d_{\mathbb{H}}(z_1, z_2) = d_{\mathbb{H}}(Qz_1, Qz_2)$ and $d_\eta(z_1, z_2) = d_{Q\eta}(Qz_1, Qz_2)$. Take an isometry Q for which $Qz_0 = i$ and $Q\eta_1 = \infty$ and denote $Q\eta_0$ by q. The geodesic $z(t)$ is mapped into the geodesic connecting i with ∞ with speed α, thus $Qz(t) = ie^{\alpha t}$. From (11), it follows that

$$d_{\eta_1}\big(z(t), z_0\big) = d_\infty\big(ie^{\alpha t}, i\big) = -\alpha t,$$
$$d_{\eta_0}\big(z(t), z_0\big) = d_q\big(ie^{\alpha t}, i\big) = \alpha t + \log\left(\frac{q^2 e^{-2\alpha t} + 1}{1 + q^2}\right) \le \alpha t.$$

From the triangle inequality, we get

$$\begin{aligned} d_{\eta_1}\big(\gamma(t), z_0\big) &\le d_{\eta_1}\big(z(t), z_0\big) + d_{\mathbb{H}}\big(\gamma(t), z(t)\big) \le -(\alpha - \varepsilon)t + b, \\ d_{\eta_0}\big(\gamma(t), z_0\big) &\le d_{\eta_0}\big(z(t), z_0\big) + d_{\mathbb{H}}\big(\gamma(t), z(t)\big) \le (\alpha + \varepsilon)t + b. \end{aligned} \tag{37}$$

The bounds in (12) now follow easily:

$$\begin{aligned} \int_0^1 e^{d_{\eta_1}(\tilde{\gamma}(t), z_0)}\, dt &\le \int_0^1 e^{-(\alpha-\varepsilon)\nu \log(\frac{1}{1-t}) + b}\, dt \\ &= e^b \int_0^1 (1-t)^{(\alpha-\varepsilon)\nu}\, dt = \frac{e^b}{1 + (\alpha - \varepsilon)\nu} < \infty \end{aligned}$$

and

$$\begin{aligned}&\int_0^1 \int_0^t e^{d_{\eta_0}(\tilde\gamma(s),z_0)+d_{\eta_1}(\tilde\gamma(t),z_0)}\,ds\,dt\\ &\quad\le e^{2b}\int_0^1\int_0^t (1-s)^{-(\alpha+\varepsilon)\nu}(1-t)^{(\alpha-\varepsilon)\nu}\,ds\,dt\\ &\quad=\frac{e^{2b}}{2(1+(\alpha-\varepsilon)\nu)(1-\varepsilon\nu)}<\infty.\end{aligned}$$ □

Recall from (14) that $\mathtt{r}\tau$ is an integral operator with kernel $K_{\mathtt{r}\tau}$. For $0<T<1$, we denote by $\mathtt{r}_T\tau$ the integral operator with kernel

$$K_{\mathtt{r}\tau}(x,y)\cdot\mathbf{1}(0\le x\le T, 0\le y\le T).$$

PROPOSITION 15 (Hilbert–Schmidt truncation). *Let $\gamma,\eta_0,\eta_1,z_0,\alpha,\nu,\tilde\gamma,z(t)$ be as in Proposition 14, and define τ according to (36).*

1. *Assume that for some $0<b$ and $0\le\kappa<1$ the following inequality holds for all $0\le t$:*

$$d_{\mathbb{H}}\big(\gamma(t),z(t)\big)\le b+t^\kappa, \tag{38}$$

and that for some $c_0<\infty$ we have

$$\alpha\nu\le c_0. \tag{39}$$

Then for any $T\in(0,1)$, we have

$$\|\mathtt{r}\tau-\mathtt{r}_T\tau\|_{\mathrm{HS}}^2\le C(1-T)^{1+\frac12\min(\alpha\nu,1)}\bigg(1+\log\frac{1}{1-T}\bigg) \tag{40}$$

with C depending only on $\eta_0,\eta_1,b,\alpha,\kappa$ and c_0 from (39).

2. *Assume that for some $0<T<1$ (38) holds for $0\le t\le\nu\log(\frac{1}{1-T})$ with some $0<b$, $0\le\kappa<1$. Assume further that $d_{\mathbb{H}}(\tilde\gamma(t),\tilde\gamma(T))\le M$ for $t\ge T$ and that (39) holds. Then we have*

$$\|\mathtt{r}\tau-\mathtt{r}_T\tau\|_{\mathrm{HS}}^2\le Ce^{2M}(1-T)^{1+\frac12\min(\alpha\nu,1)}\bigg(1+\log\frac{1}{1-T}\bigg), \tag{41}$$

with C depending only on $\eta_0,\eta_1,b,\alpha,\kappa$ and c_0 from (39).

PROOF. We denote the representation of γ in the half-plane by $x+iy$ and use $\tilde x+i\tilde y$ for the representation of $\tilde\gamma$. We represent η_0,η_1 with nonzero vectors u_0,u_1 that satisfy $u_0^tJu_1=1$. Recall the integral kernel of $\mathtt{r}\tau$ from (14). From the definition of $\mathtt{r}_T\tau$, we get

$$2\|\mathtt{r}\tau-\mathtt{r}_T\tau\|_{\mathrm{HS}}^2=\int_T^1\int_s^1|a(s)|^2|c(t)|^2\,dt\,ds+\int_0^T\int_T^1|a(s)|^2|c(t)|^2\,dt\,ds,$$

where $a(s)=X(s)u_0$ and $c(s)=X(s)u_1$ with $X=\frac{1}{\sqrt{\tilde y}}\big(\begin{smallmatrix}1&-\tilde x\\0&\tilde y\end{smallmatrix}\big)$.

From (11), one can check that if $u\in\mathbb{R}^2$ is a nonzero vector then $e^{d_u(x+iy,i)}=\frac{|Xu|^2}{|u|^2}$. Using the triangle inequality, we get

$$\begin{aligned}|a(s)|^2&=|X(s)u_0|^2=|u_0|^2e^{d_{\eta_0}(\tilde\gamma,i)}\le|u_0|^2e^{d_{\eta_0}(\tilde\gamma,z_0)+d_{\mathbb{H}}(z_0,i)},\\ |c(s)|^2&=|X(s)u_1|^2=|u_1|^2e^{d_{\eta_1}(\tilde\gamma,i)}\le|u_1|^2e^{d_{\eta_1}(\tilde\gamma,z_0)+d_{\mathbb{H}}(z_0,i)}.\end{aligned} \tag{42}$$

Using now (37) and recalling $\tilde\gamma(t)=\gamma(\nu\log(\frac{1}{1-t}))$, we get that if (38) holds for all t then

$$|c(t)|^2\le C_0(1-t)^{\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})},\qquad |a(t)|^2\le C_0(1-t)^{-\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})},\tag{43}$$

where C_0 depends on u_0,u_1 and b. This leads to

$$\begin{aligned}&2\|\mathtt{r}\tau-\mathtt{r}_T\tau\|_{\mathrm{HS}}^2\\ &\quad\le C_0^2\int_T^1\int_s^1(1-t)^{\alpha\nu}(1-s)^{-\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})}e^{\nu^\kappa\log^\kappa(\frac{1}{1-s})}\,dt\,ds\\ &\qquad+C_0^2\int_0^T(1-s)^{-\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-s})}\,ds\int_T^1(1-t)^{\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})}\,dt.\end{aligned}\tag{44}$$

Note that for $0<s<1$ and $0<\varepsilon<1$ we have

$$\begin{aligned}\int_s^1(1-t)^{\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})}\,dt&\le\int_s^1(1-t)^{(1-\varepsilon)\alpha\nu}\,dt\max_{0\le t\le 1}(1-t)^{\varepsilon\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})}\\ &\le(1-s)^{(1-\varepsilon)\alpha\nu+1}C(\varepsilon\alpha,\kappa),\end{aligned}$$

where $C(\varepsilon\alpha,\kappa)=\max_{0\le x}e^{-\varepsilon\alpha x+x^\kappa}$ is a positive constant depending on $\varepsilon\alpha$ and κ. This leads to

$$\begin{aligned}&\int_T^1\int_s^1(1-s)^{-\alpha\nu}e^{\nu^\kappa\log(\frac{1}{1-s})}(1-t)^{\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})}\,dt\,ds\\ &\quad\le C(\varepsilon\alpha,\kappa)\int_T^1(1-s)^{1-\varepsilon\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-s})}\,ds\le C(\varepsilon\alpha,\kappa)^2(1-T)^{2-2\varepsilon\alpha\nu},\end{aligned}$$

where for the validity of the last step we also assume $2\varepsilon c_0\le 1$. Moreover,

$$\begin{aligned}&\int_0^T(1-s)^{-\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-s})}\,ds\int_T^1(1-t)^{\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-t})}\,dt\\ &\quad\le C(\varepsilon\alpha,\kappa)^2(1-T)^{1+(1-\varepsilon)\alpha\nu}\int_0^T(1-s)^{-(1+\varepsilon)\alpha\nu}\,ds\\ &\quad\le C(\varepsilon\alpha,\kappa)^2\log\frac{1}{1-T}\cdot(1-T)^{1+\min(1-2\varepsilon\alpha\nu,(1-\varepsilon)\alpha\nu)},\end{aligned}$$

where the last bound follows from the inequality

$$\int_0^T(1-s)^r\,ds\le\log\frac{1}{1-T}\cdot(1-T)^{\min(1+r,0)},\tag{45}$$

which holds for all $0\le T<1$ and $r\in\mathbb{R}$.

Now choose $\varepsilon=\min(\frac12,\frac{1}{4c_0})$ with c_0 from (39). Collecting all of our bounds and returning to (44), we get the estimate in (40).

Now assume that (38) holds for $0\le t\le\nu\log(\frac{1}{1-T})$, and $d_{\mathbb{H}}(\tilde\gamma(t),\tilde\gamma(T))\le M$ for $t\ge T$. Then for $0\le t\le T$, we still have (43), while for $t\ge T$ we can use $d_\eta(x,y)\le d_\eta(x,z)+d_{\mathbb{H}}(y,z)$ together with (42) to get

$$\begin{aligned}|c(t)|^2&\le C_0e^M(1-T)^{\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-T})},\\ |a(t)|^2&\le C_0e^M(1-T)^{-\alpha\nu}e^{\nu^\kappa\log^\kappa(\frac{1}{1-T})},\end{aligned}$$

for $T \le t < 1$. This gives

$$
\begin{aligned}
&2\|\mathtt{r}\tau - \mathtt{r}_T\tau\|_{\mathrm{HS}}^2\\
&\quad \le C_0^2 e^{2M} \int_T^1 \int_s^1 e^{2\nu^\kappa \log^\kappa(\frac{1}{1-T})}\,dt\,ds\\
&\qquad + C_0^2 e^M \int_0^T (1-s)^{-\alpha\nu} e^{\nu^\kappa \log^\kappa(\frac{1}{1-s})}\,ds \int_T^1 (1-T)^{\alpha\nu} e^{\nu^\kappa \log^\kappa(\frac{1}{1-T})}\,dt.
\end{aligned}
$$

The bound (41) now follows by using similar estimates as in the proof of (40) from (44). □

PROPOSITION 16 (Hilbert–Schmidt approximation). *Let $\gamma, \eta_0, \eta_1, z_0, \alpha, \nu, \tilde\gamma, z(t)$ be as in Proposition 14, and define τ according to (36). Assume that (38) holds for $0 \le t \le \nu\log(\frac{1}{1-T})$ for some $T \in (0,1)$, $0 < b$ and $0 \le \kappa < 1$.*

Suppose that the path $\tilde\gamma_1$ is measurable and for $0 \le t \le \nu\log(\frac{1}{1-T})$ we have

$$
\sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde\gamma(t), \tilde\gamma_1(t))\right)^2 \le \min(\delta(1-t)^{-1}, M) \tag{46}
$$

for some $M, \delta > 0$.

Consider $\tau_1 = \mathtt{Dir}(\{\tilde\gamma_1(t), t \ge 0\}, \eta_0, \eta_1)$, define $\mathtt{r}_T\tau$ as in Proposition 15, and $\mathtt{r}_T\tau_1$ similarly. Then

$$
\|\mathtt{r}_T\tau - \mathtt{r}_T\tau_1\|_{\mathrm{HS}}^2 \le C(M+1)\delta, \tag{47}
$$

with a constant C depending only on $\eta_0, \eta_1, \kappa, b, \alpha$ and ν.

PROOF. Denote the representation of $\tilde\gamma_1$ in the half-plane by $\tilde x_1 + i\tilde y_1$.

The hyperbolic distance formula (10) in the upper half-plane representation gives

$$
4\sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde\gamma, \tilde\gamma_1)\right)^2 = \frac{(\tilde x - \tilde x_1)^2}{\tilde y \tilde y_1} + \left(\sqrt{\frac{\tilde y}{\tilde y_1}} - \sqrt{\frac{\tilde y_1}{\tilde y}}\right)^2.
$$

Consider a, c, defined as in the proof of Proposition 15 and the analogously defined a_1, c_1.

If $u \in \mathbb{R}^2$ is a nonzero vector, $X = \frac{1}{\sqrt{\tilde y}}\left(\begin{smallmatrix} 1 & -\tilde x \\ 0 & \tilde y \end{smallmatrix}\right)$ and $X_1 = \frac{1}{\sqrt{\tilde y_1}}\left(\begin{smallmatrix} 1 & -\tilde x_1 \\ 0 & \tilde y_1 \end{smallmatrix}\right)$ then

$$
\frac{|Xu - X_1 u|}{|Xu|} = \frac{|(I - X_1 X^{-1})Xu|}{|Xu|} \le \|I - X_1 X^{-1}\|_2.
$$

An explicit computation gives

$$
\begin{aligned}
\|I - X_1 X^{-1}\|_2^2 &= \left(\frac{x - x_1}{\sqrt{\tilde y \tilde y_1}}\right)^2 + \left(1 - \sqrt{\frac{\tilde y_1}{\tilde y}}\right)^2 + \left(1 - \sqrt{\frac{\tilde y}{\tilde y_1}}\right)^2\\
&\le \left(\frac{\tilde x - \tilde x_1}{\sqrt{\tilde y \tilde y_1}}\right)^2 + \left(\sqrt{\frac{\tilde y}{\tilde y_1}} - \sqrt{\frac{\tilde y_1}{\tilde y}}\right)^2\\
&= 4\sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde x + i\tilde y, \tilde x_1 + i\tilde y_1)\right)^2.
\end{aligned}
$$

This yields

$$
\frac{|c - c_1|^2}{|c|^2} \le 4\sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde\gamma, \tilde\gamma_1)\right)^2, \qquad \frac{|a - a_1|^2}{|a|^2} \le 4\sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde\gamma, \tilde\gamma_1)\right)^2.
$$

To compute $\|\mathtt{r}_T\tau - \mathtt{r}_T\tau_1\|_{\mathrm{HS}}^2$, we need to estimate $\mathrm{tr}(\Delta K(s,t)\Delta K(s,t)^t)$ where $\Delta K(s,t) = \frac{1}{2}(a(s)c(t)^t - a_1(s)c_1(t)^t)$. Using the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
&4\,\mathrm{tr}\big(\Delta K(s,t)\Delta K(s,t)^t\big)\\
&\quad\le 3\big(|a(s)|^2|c(t)-c_1(t)|^2 + |a(s)-a_1(s)|^2|c(t)|^2\\
&\qquad + |a(s)-a_1(s)|^2|c(t)-c_1(t)|^2\big).
\end{aligned}$$

The previous estimates with (46) yield

$$\begin{aligned}
&\|\mathtt{r}_T\tau - \mathtt{r}_T\tau_1\|_{\mathrm{HS}}^2\\
&\quad= 2\int_0^T\int_0^t \mathrm{tr}\big(\Delta K(s,t)\Delta K(s,t)^t\big)\,ds\,dt\\
&\quad\le 24\int_0^T\int_0^t |a(s)|^2|c(t)|^2\biggl(\sinh\biggl(\frac{d_{\mathbb{H}}(s)}{2}\biggr)^2 + \sinh\biggl(\frac{d_{\mathbb{H}}(t)}{2}\biggr)^2\\
&\qquad + \sinh\biggl(\frac{d_{\mathbb{H}}(s)}{2}\biggr)^2\sinh\biggl(\frac{d_{\mathbb{H}}(t)}{2}\biggr)^2\biggr)\,ds\,dt\\
&\quad\le 24\int_0^T\int_0^t |a(s)|^2|c(t)|^2\biggl(\delta(1-t)^{-1} + \delta(1-s)^{-1}\\
&\qquad + \frac{1}{2}\delta M\big((1-t)^{-1} + (1-s)^{-1}\big)\biggr)\,ds\,dt,
\end{aligned}$$

where we used the notation $d_{\mathbb{H}}(t) = d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}_1(t))$. Using the arguments in the proof of Proposition 15, we get that (43) holds for $0 \le t \le T$, with a constant C_1 depending on η_0, η_1 and b. Using the temporary notation $g(t) = \nu^\kappa \log^\kappa(\frac{1}{1-t})$, this leads to

$$\begin{aligned}
&\|\mathtt{r}_T\tau - \mathtt{r}_T\tau_1\|_{\mathrm{HS}}^2\\
&\quad\le 24C_1^2\delta\biggl(1+\frac{M}{2}\biggr)\int_0^T\int_0^t e^{g(t)+g(s)}\big((1-s)^{-\alpha\nu-1}(1-t)^{\alpha\nu}\\
&\qquad + (1-s)^{-\alpha\nu}(1-t)^{\alpha\nu-1}\big)\,ds\,dt\\
&\quad\le 24C_1^2\delta(2+M)\int_0^1 e^{2g(t)}(1-t)^{\alpha\nu-1}\int_0^t(1-s)^{-\alpha\nu}\,ds\,dt\\
&\quad\le 24C_1^2\delta(2+M)\int_0^1 e^{2g(t)}\log\biggl(\frac{1}{1-t}\biggr)(1-t)^{\min(\alpha\nu-1,0)}\,dt,
\end{aligned}$$

where we used (45) in the last step.

Since the integral $\int_0^1 e^{2\nu^\kappa\log^\kappa(\frac{1}{1-t})}\log(\frac{1}{1-t})(1-t)^{\min(\alpha\nu-1,0)}\,dt$ is finite for any given $0 \le \kappa < 1$, $\nu > 0$ and $\alpha > 0$, the bound (47) now follows. □

5. Proof of Theorem 1. We now return to the proof of our main theorem.

PROOF OF THEOREM 1. Let $\mathcal{B}$ be a hyperbolic Brownian motion. Set $\tilde{\mathcal{B}}(t) = \mathcal{B}(-\frac{4}{\beta}\log(1-t))$ for $0 \le t < 1$. Set $\eta_0 = \infty$, $\eta_1 = \mathcal{B}(\infty)$ and set $\mathtt{Sine}_\beta = \mathtt{Dir}(\tilde{\mathcal{B}}, \eta_0, \eta_1)$.

Let $U_k, \xi_k, k \ge 1$ be random variables independent of each other and $\mathcal{B}$ with distributions given in Proposition 12, and let $\tau_{n,k}$ be the stopping times constructed there. According to the proposition, the path $\tilde{\mathcal{B}}_n(t) = \mathcal{B}(\tau_{n,\lceil(1-t)n\rceil})$, $t \in [0,1)$ has the same distribution as the path in the construction of $\mathtt{Circ}_{\beta,n}$ in Theorem 8, and we may write $\mathtt{Circ}_{\beta,n} = \mathtt{Dir}(\tilde{\mathcal{B}}_n, \eta_0, \eta_1)$.

To prove the bound (1), we set $T_n = 1 - \frac{\log^6 n}{n}$. Recall the definition of $\mathtt{r}_T$ from Section 4 and write

$$\begin{aligned}
&\|\mathtt{rSine}_\beta - \mathtt{rCirc}_{\beta,n}\|_{\mathrm{HS}}^2 \\
&\quad \le 3\|\mathtt{r}_{T_n}\mathtt{Sine}_\beta - \mathtt{r}_{T_n}\mathtt{Circ}_{\beta,n}\|_{\mathrm{HS}}^2 \\
&\qquad + 3\|\mathtt{rSine}_\beta - \mathtt{r}_{T_n}\mathtt{Sine}_\beta\|_{\mathrm{HS}}^2 \\
&\qquad + 3\|\mathtt{rCirc}_{\beta,n} - \mathtt{r}_{T_n}\mathtt{Circ}_{\beta,n}\|_{\mathrm{HS}}^2.
\end{aligned} \tag{48}$$

We will use Propositions 13, 15 and 16 to estimate the three terms on the right. Let $z(t)$ be the point moving with speed $1/2$ on the geodesic connecting $\mathcal{B}(0)$ to $\mathcal{B}(\infty)$. From Lemma 21 of the Appendix, it follows that for $\kappa = 2/3$ there is a random b so that a.s. for all $t \ge 0$ we have

$$d_{\mathbb{H}}\big(\mathcal{B}(t), z(t)\big) \le b + t^{2/3}. \tag{49}$$

Applying the first statement of Proposition 15 for $\mathtt{Sine}_\beta = \mathtt{Dir}(\tilde{\mathcal{B}}(t), \eta_0, \eta_1)$ with $\gamma = \mathcal{B}$, $\nu = \frac{4}{\beta}$, $\alpha = 1/2$, $\kappa = 2/3$ and $c_0 = \frac{2}{\beta}$, we get

$$\|\mathtt{rSine}_\beta - \mathtt{r}_{T_n}\mathtt{Sine}_\beta\|_{\mathrm{HS}}^2 \le C\left(\frac{\log^6 n}{n}\right)^{1+\frac{1}{2}\min(\frac{2}{\beta},1)}\left(1 + \log\frac{n}{\log^6 n}\right)$$

with a random C depending on $\mathcal{B}$ and β.

Recall from Proposition 13 that for $n \ge N^*$ we have

$$d_{\mathbb{H}}\big(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t)\big) \le \frac{\log^{3-1/8} n}{\sqrt{(1-t)n}} \le 1 \quad \text{if } 0 \le t \le T_n, \tag{50}$$

$$d_{\mathbb{H}}\big(\tilde{\mathcal{B}}_n(T_n), \tilde{\mathcal{B}}_n(t)\big) \le (\log\log n)^4 \quad \text{if } T_n \le t < 1. \tag{51}$$

Let $\mathcal{B}_n(t) = \tilde{\mathcal{B}}_n(1 - e^{-\frac{\beta}{4}t})$, then $\tilde{\mathcal{B}}_n(t) = \mathcal{B}_n(\frac{4}{\beta}\log(\frac{1}{1-t}))$ for $0 \le t < 1$. From (49), (50) and the triangle inequality, we get

$$d_{\mathbb{H}}\big(\mathcal{B}_n(t), z(t)\big) \le b + 1 + t^{2/3} \quad \text{for } 0 \le t \le \frac{4}{\beta}\log\left(\frac{1}{1-T_n}\right).$$

Recall that $\mathtt{Circ}_{\beta,n} = \mathtt{Dir}(\tilde{\mathcal{B}}_n(t), \eta_0, \eta_1)$. Applying the second statement of Proposition 15 with $\gamma = \mathcal{B}_n$, $\nu = \frac{4}{\beta}$, $\alpha = 1/2$, $\kappa = 2/3$, $c_0 = \frac{2}{\beta}$ and $M = (\log\log n)^4$, $T = T_n$ we get

$$\begin{aligned}
&\|\mathtt{rCirc}_{\beta,n} - \mathtt{r}_{T_n}\mathtt{Circ}_{\beta,n}\|_{\mathrm{HS}}^2 \\
&\quad \le C e^{2(\log\log n)^4}\left(\frac{\log^6 n}{n}\right)^{1+\frac{1}{2}\min(\frac{2}{\beta},1)}\left(1 + \log\frac{n}{\log^6 n}\right)
\end{aligned}$$

if $n \ge N^*$, with a random C depending on $\mathcal{B}$ and β.

Since $\sinh(x/2)^2 \le x^2$ for $0 \le x \le 1$, from (50) we get

$$\sinh\left(\frac{1}{2}d_{\mathbb{H}}\big(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t)\big)\right)^2 \le \frac{\log^{6-1/4} n}{n}(1-t)^{-1} \le 1 \quad \text{if } 0 \le t \le T_n.$$

Hence we may use Proposition 16 with $\gamma = \mathcal{B}$, $\gamma_1 = \mathcal{B}_n$, $\nu = \frac{4}{\beta}$, $\alpha = 1/2$, $T = T_n$, $\delta = \frac{\log^{6-1/4} n}{n}$, $M = 1$, $\kappa = 2/3$ to get

$$\|\mathtt{r}_{T_n}\mathtt{Sine}_\beta - \mathtt{r}_{T_n}\mathtt{Circ}_{\beta,n}\|_{\mathrm{HS}}^2 \le C\frac{\log^{6-1/4} n}{n} \quad \text{if } n \ge N^*,$$

with a random C depending on $\mathcal{B}$ and β. Collecting our estimates, going back to (48) and modifying the random lower bound N^* appropriately, we get

$$\|\mathtt{rSine}_\beta - \mathtt{rCirc}_{\beta,n}\|_{\mathrm{HS}}^2 \le \frac{\log^6 n}{n} \quad \text{for } n \ge N^*.$$

This completes the proof of Theorem 1. □

Before proving the three statements of Corollary 2, we state a law of large numbers for the points of Sine_β.

PROPOSITION 17. *Suppose that the points of* Sine_β *are given by* $\lambda_k, k \in \mathbb{Z}$ *with* $\lambda_0 < 0 \le \lambda_1$. *Then with probability one, we have*

$$\lim_{k\to\infty} \frac{\lambda_k}{k} = \lim_{k\to-\infty} \frac{\lambda_k}{k} = 2\pi. \tag{52}$$

PROOF. In Holcomb and Valkó (2015), it was shown that $\frac{1}{\lambda}\#\{\lambda_k : 0 \le \lambda_k \le \lambda\}$ satisfies a large deviation principle as $\lambda \to \infty$ with scale λ^2 and rate function $\beta I(\rho)$, where $I(\frac{1}{2\pi}) = 0$ is the global minimum. From this, the statement follows for $k \to \infty$ by a simple Borel–Cantelli argument. The $k \to -\infty$ case follows similarly using the symmetry of the Sine_β process. □

PROOF OF COROLLARY 2. The bound (2) follows directly from (1) and the Hoffman–Wielandt inequality for compact integral operators (see, e.g., Bhatia and Elsner (1994)). Note that it might be possible to get a sharper bound by comparing the eigenfunctions as well.

From Proposition 17, we see that there is a random constant C so that a.s. $|\lambda_k| \le C(|k|+1)$ for all k. Set $a_n = \frac{n^{1/4}}{\log^2 n}$. From (1), it follows that for large enough n we have

$$\sup_k |\lambda_k^{-1} - \lambda_{k,n}^{-1}| \le \sqrt{\sum_k |\lambda_k^{-1} - \lambda_{k,n}^{-1}|^2} \le \frac{\log^3 n}{n^{1/2}}. \tag{53}$$

Let $b_{k,n} = \lambda_k^{-1} - \lambda_{k,n}^{-1}$, then

$$(\lambda_{k,n} - \lambda_k)(1 - b_{k,n}\lambda_k) = b_{k,n}\lambda_k^2. \tag{54}$$

For large enough n, we have

$$\max_{|k|\le a_n} |b_{k,n}\lambda_k| \le \frac{\log^3 n}{n^{1/2}} C(a_n+1) = C\left(\frac{\log n}{n^{1/4}} + \frac{\log^3 n}{n^{1/2}}\right),$$

which means that $\lim_{n\to\infty} \max_{|k|\le a_n} |b_{k,n}\lambda_k| = 0$ a.s. From (54), for large enough n, we have

$$\max_{|k|\le a_n} |\lambda_{k,n} - \lambda_k| \le 2C^2(a_n+1)^2 \max_{|k|\le a_n} |b_{k,n}| \le 4C^2 a_n^2 \frac{\log^3 n}{n^{1/2}} = \frac{4C^2}{\log n}.$$

This completes the proof of (3).

By Proposition 17, we may choose a random $C > 0$ so that $|\lambda_k| \le C\sqrt{1+k^2}$ for all k. Then we have $\frac{1}{C\sqrt{1+k^2}} \le \frac{1}{|\lambda_k|}$ and for $n \ge N$, $|k| \le n^{1/2-\varepsilon}$ we have

$$\frac{1}{|\lambda_{k,n}|} \ge \frac{1}{|\lambda_k|} - \left|\frac{1}{\lambda_k} - \frac{1}{\lambda_{k,n}}\right| \ge \frac{1}{C\sqrt{1+k^2}} - \frac{\log^3 n}{\sqrt{n}} \ge \frac{1}{2C\sqrt{1+k^2}}.$$

(For the last inequality, one might need to change N.) Then for such k and n, we get

$$|\lambda_k - \lambda_{k,n}| = \left|\frac{1}{\lambda_k} - \frac{1}{\lambda_{k,n}}\right| |\lambda_k||\lambda_{k,n}| \le \frac{\log^3 n}{n^{1/2}} 2C^2(1+k^2) \le \frac{1+k^2}{n^{1/2-\varepsilon}},$$

again by setting a large enough random lower bound on n. □

6. Beta dependence in the $\mathtt{Sine}_\beta$ operator. The construction of the $\mathtt{Sine}_\beta$ operator provides a natural coupling of this operator for all values of β. The techniques we developed for the proof of Theorem 1 can be used to estimate how the operator $\mathtt{Sine}_\beta$ (and the process Sine_β) depends on the value of β in this coupling.

PROOF OF THEOREM 3. We start with the case $\beta' < \infty$. We have

$$\mathtt{Sine}_\beta = \mathtt{Dir}(\tilde\gamma, \eta_0, \eta_1), \qquad \mathtt{Sine}_{\beta'} = \mathtt{Dir}(\tilde\gamma_1, \eta_0, \eta_1),$$

where $\tilde\gamma(t) = \mathcal{B}(\frac{4}{\beta}\log(\frac{1}{1-t}))$, $\tilde\gamma_1(t) = \mathcal{B}(\frac{4}{\beta'}\log(\frac{1}{1-t}))$, $\eta_0 = \infty$ and $\eta_1 = \mathcal{B}(\infty)$. We will estimate $\|\mathtt{r}\mathtt{Sine}_\beta - \mathtt{r}\mathtt{Sine}_{\beta'}\|_{\mathrm{HS}}$ using Propositions 15 and 16. Note that with probability one $\eta_0 \neq \eta_1$, and hence a.s. 0 is not an eigenvalue for any of the $\mathtt{Sine}_\beta$ operators.

Set $T = 1-\delta < 1$. Cutting off $\mathtt{Sine}_\beta$ and $\mathtt{Sine}_{\beta'}$ at T and using the first statement of Proposition 15 with $\alpha = 1/2$, $\kappa = 2/3$, $c_0 = \frac{2}{\theta}$, and $\nu = \frac{4}{\beta}$ and $\frac{4}{\beta'}$, respectively, as in the proof of Theorem 1) gives

$$\|\mathtt{r}\mathtt{Sine}_\beta - \mathtt{r}_T\mathtt{Sine}_\beta\|_{\mathrm{HS}}^2 \le C_0\delta^{1+\frac12\min(\frac{2}{\beta},1)}(1+\log\delta^{-1}),$$

$$\|\mathtt{r}\mathtt{Sine}_{\beta'} - \mathtt{r}_T\mathtt{Sine}_{\beta'}\|_{\mathrm{HS}}^2 \le C_0\delta^{1+\frac12\min(\frac{2}{\beta'},1)}(1+\log\delta^{-1})$$

with a random C_0 depending only on $\mathcal{B}$.

To estimate $\|\mathtt{r}_T\mathtt{Sine}_\beta - \mathtt{r}_T\mathtt{Sine}_{\beta'}\|_{\mathrm{HS}}^2$, we first bound $d_{\mathbb{H}}(\tilde\gamma(t), \tilde\gamma_1(t))$ in $[0,T]$. For $0 < t \le T$, we have $(\frac{4}{\beta'} - \frac{4}{\beta})\log(\frac{1}{1-t}) \le \delta\log(\delta^{-1}) < 1$ and by Proposition 22 (and the comment following the proposition), we get the bound

$$\begin{aligned} &d_{\mathbb{H}}\left(\mathcal{B}\left(\frac{4}{\beta}\log\left(\frac{1}{1-t}\right)\right), \mathcal{B}\left(\frac{4}{\beta'}\log\left(\frac{1}{1-t}\right)\right)\right)^2 \\ &\quad < C^2\delta\log\left(\frac{1}{1-t}\right)\log\left(2 + \frac{\frac{4}{\beta'}\log(\frac{1}{1-t})+1}{\delta\log(\frac{1}{1-t})}\right) \end{aligned} \tag{55}$$

for $0 \le t \le T$ with a random C depending only on $\mathcal{B}$.

Since $0 \le t \le T = 1-\delta$, the right-hand side of (55) can be bounded as

$$C^2\delta\log\left(\frac{1}{1-t}\right)\left(\log\left(2+\frac{4}{\theta}\delta^{-1}\right) + \log\left(1+\frac{1}{\delta\log(\frac{1}{1-t})}\right)\right) \le C^2 c_1,$$

with a constant c_1 depending only on θ.

Using $\log(\frac{1}{1-t}) \le (1-t)^{-1/2}$ (which holds for $0 \le t < 1$), we also get the bound

$$d_{\mathbb{H}}(\tilde\gamma(t), \tilde\gamma_1(t))^2 < \min\left(C_1\delta\log\left(\frac{1}{\delta}\right)(1-t)^{-1/2}, C_1\right), \qquad t \in [0,T]$$

with a constant C_1 depending on $\mathcal{B}$ and θ. We can turn this into an upper bound of the form

$$\begin{aligned} &\sinh(\frac12 d_{\mathbb{H}}(\tilde\gamma(t), \tilde\gamma_1(t))^2 \\ &\quad \le \min\left(C_2\delta\log\left(\frac{1}{\delta}\right)(1-t)^{-1/2}, C_2\right), \qquad t \in [0,T], \end{aligned} \tag{56}$$

with C_2 depending on $\mathcal{B}$ and θ.

Using the arguments in the proof of Proposition 16, we get the bounds

$$\begin{aligned} &\|\mathtt{r}_T\mathtt{Sine}_\beta - \mathtt{r}_T\mathtt{Sine}_{\beta'}\|_{\mathrm{HS}}^2 \\ &\quad \le C_3\delta\log\left(\frac{1}{\delta}\right)\int_0^T\int_0^t e^{g(t)+g(s)}((1-s)^{-\frac{2}{\beta}-1/2}(1-t)^{\frac{2}{\beta}} \end{aligned}$$

$$
\begin{aligned}
&+ (1-s)^{-\frac{2}{\beta}}(1-t)^{\frac{2}{\beta}-1/2}\big)\,ds\,dt \\
&\le 2C_3\delta\log\left(\frac{1}{\delta}\right)\int_0^1 e^{2g(t)}(1-t)^{\frac{2}{\beta}-1/2}\int_0^t (1-s)^{-\frac{2}{\beta}}\,ds\,dt \\
&\le 2C_3\delta\log\left(\frac{1}{\delta}\right)\int_0^1 e^{2g(t)}\log\left(\frac{1}{1-t}\right)(1-t)^{\min(\frac{2}{\beta}-1/2,1/2)}\,dt,
\end{aligned}
$$

with $g(t) = (4/\beta)^{2/3}\log^{2/3}(\frac{1}{1-t})$ and C_3 still only depending on $\mathcal{B}$ and θ. Since $\theta \le \beta$, we have

$$
\begin{aligned}
&\int_0^1 e^{2g(t)}\log\left(\frac{1}{1-t}\right)(1-t)^{\min(\frac{2}{\beta}-1/2,1/2)}\,dt \\
&\quad\le \int_0^1 e^{2g(t)}\log\left(\frac{1}{1-t}\right)(1-t)^{\frac{1}{\beta}\min(2,\theta)-1/2}\,dt \\
&\quad\le \int_0^1 \log\left(\frac{1}{1-t}\right)(1-t)^{-1/2}\,dt\,\max_{x>0} e^{2x^{2/3}-\frac{1}{4}\min(2,\theta)x} \le C_4,
\end{aligned}
$$

with C_4 depending only on θ. This gives

$$
\|\mathtt{r}_T\mathtt{Sine}_\beta - \mathtt{r}_T\mathtt{Sine}_{\beta'}\|_{\mathrm{HS}}^2 \le C_5\delta\log\left(\frac{1}{\delta}\right),
$$

with C_5 depending only on $\mathcal{B}$ and θ. (Note that we needed the bound (56) that was slightly better than the assumption (46) in Proposition 16 to get an upper bound here that does not depend on β, β'.)

Collecting all the terms gives

$$
\|\mathtt{rSine}_\beta - \mathtt{rSine}_{\beta'}\|_{\mathrm{HS}}^2 \le C\delta\log\left(\frac{1}{\delta}\right),
$$

with a C depending only on $\mathcal{B}$ and θ. The Hoffman–Wielandt inequality completes the proof of (5).

To treat the $\beta' = \infty$ case, we note that here $\tilde{\gamma}_1(t) = \mathcal{B}(0)$, hence the integral kernel of $\mathtt{rSine}_{\beta'}$ is constant on the sets $\{(s,t) : 0 \le s < t < 1\}$ and $\{(s,t) : 0 \le t \le s < 1\}$, with the constants depending only on $\eta_1 = \mathcal{B}(\infty)$. Setting $\delta = \frac{4}{\beta} \le 1/3$ and $T = 1-\delta$, we have

$$
\|\mathtt{rSine}_\infty - \mathtt{r}_T\mathtt{Sine}_\infty\|_{\mathrm{HS}}^2 \le C_0\delta
$$

with C_0 depending only $\mathcal{B}(\infty)$.

Using (56), we have

$$
\|\mathtt{rSine}_\beta - \mathtt{r}_T\mathtt{Sine}_\beta\|_{\mathrm{HS}}^2 \le C_1\delta\log\left(\frac{1}{\delta}\right)
$$

with C_1 depending only on $\mathcal{B}$ and θ.

The term $\|\mathtt{r}_T\mathtt{Sine}_\beta - \mathtt{r}_T\mathtt{Sine}_\infty\|_{\mathrm{HS}}^2$ can be bounded similarly as in the case $\beta' < \infty$, leading to the same upper bound. This completes the proof of (5) in the $\beta' = \infty$ case. $\square$

APPENDIX

In the first part of the Appendix, we collect the proofs of Lemmas 9 and 11 that were used in the single step coupling of Proposition 10. In the second part, we collect some estimates on the hyperbolic Brownian motion.

A.1. Proof of the coupling statements. We now return to the proof of Lemma 9. Let $\mathcal{B}$ be a hyperbolic Brownian motion and set $\zeta = \zeta_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$.

The circumference of a hyperbolic circle of radius r is $2\pi \sinh(r)$ and the density function of the hyperbolic BM at time t and distance r is given by the following formula (see, e.g., Karpelevič, Tutubalin and Šur (1959)):

$$g(r,t) = \frac{\sqrt{2}e^{-t/8}}{(2\pi t)^{3/2}} \int_r^\infty \frac{s e^{-\frac{s^2}{2t}}}{\sqrt{\cosh s - \cosh r}}\, ds.$$

From this, the density of ζ_t is

$$p_\zeta(r,t) = 2\pi \sinh(r) g(r,t) = \frac{e^{-t/8}\sinh(r)}{\sqrt{\pi} t^{3/2}} \int_r^\infty \frac{s e^{-\frac{s^2}{2t}}}{\sqrt{\cosh s - \cosh r}}\, ds, \tag{57}$$

and the tail of the cumulative distribution function is

$$\begin{aligned} 1 - F_\zeta(r) &= \int_r^\infty \sinh(u) \frac{e^{-t/8}}{\sqrt{\pi} t^{3/2}} \int_u^\infty \frac{s e^{-\frac{s^2}{2t}}}{\sqrt{\cosh s - \cosh u}}\, ds\, du \\ &= \int_r^\infty \frac{e^{-t/8} s e^{-\frac{s^2}{2t}}}{\sqrt{\pi} t^{3/2}} \int_r^s \frac{\sinh(u)}{\sqrt{\cosh s - \cosh u}}\, du\, ds \\ &= \int_r^\infty \frac{2e^{-t/8} s e^{-\frac{s^2}{2t}}}{\sqrt{\pi} t^{3/2}} \sqrt{\cosh(s) - \cosh(r)}\, ds. \end{aligned} \tag{58}$$

Let $Y = Y_\gamma = \log(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}})$ where ξ has distribution $\mathrm{Beta}(1,\gamma)$. We record the cumulative distribution function F_Y and the probability density function p_Y of Y, which follow from direct computation with the Beta distribution:

$$F_Y(r) = 1 - \operatorname{sech}^{2\gamma}\left(\frac{r}{2}\right), \qquad p_Y(r) = \gamma \sinh\left(\frac{r}{2}\right) \operatorname{sech}^{2\gamma+1}\left(\frac{r}{2}\right). \tag{59}$$

We start with a simple estimate.

LEMMA 18. *For $0 \le r \le s$, we have*

$$\frac{1}{\sqrt{2}}\sqrt{s^2 - r^2} \le \sqrt{\cosh(s) - \cosh(r)} \le \frac{1}{\sqrt{2}}\sqrt{s^2 - r^2} \exp\left(\frac{r^2 + s^2}{24}\right). \tag{60}$$

PROOF. The statement follows from the bound $1 \le \frac{\sinh(x)}{x} \le e^{x^2/6}$ and

$$\frac{2(\cosh(s) - \cosh(r))}{s^2 - r^2} = \frac{\sinh((s-r)/2)}{(s-r)/2} \cdot \frac{\sinh((s+r)/2)}{(s+r)/2}.$$ □

Applying Lemma 18 to (57) and (58) and computing the resulting integrals directly, we get the following bounds:

$$p_\zeta(r,t) \ge \left(1 + \frac{t}{12}\right)^{-1/2} \frac{1}{t} e^{-\frac{r^2}{2t} - \frac{r^2}{12} - \frac{t}{8}} \sinh(r) =: p_-(r,t), \tag{61}$$

$$p_\zeta(r,t) \le \frac{1}{t} e^{-\frac{r^2}{2t} - \frac{t}{8}} \sinh(r) =: p_+(r,t), \tag{62}$$

$$1 - F_\zeta(r) \le \left(1 - \frac{t}{12}\right)^{-3/2} e^{-\frac{r^2}{2t} + \frac{r^2}{12} - \frac{t}{8}}, \tag{63}$$

where the last bound is valid for $0 < t < 12$.

We also record the following bounds on $\log\cosh(r)$ which can be readily checked by differentiation and Taylor expansion:

$$\log\cosh(x) \le \frac{x^2}{2}, \quad \text{for all } x, \tag{64}$$

$$\log\cosh(x) \ge \frac{x^2}{2} - \frac{x^4}{12}, \quad \text{for } 0 \le x \le 1. \tag{65}$$

PROOF OF (A) IN LEMMA 9. From (59) and (64), we get

$$1 - F_Y(r) = \exp(-2\gamma \log\cosh(r/2)) \ge \exp\left(-\gamma \frac{r^2}{4}\right) = \exp\left(-\frac{r^2}{2t} + \frac{r^2}{8}\right).$$

One can check that if $0 < t \le 1$ and $t \le r$ then

$$\left(1 - \frac{t}{12}\right)^{-3/2} e^{-\frac{r^2}{2t} + \frac{r^2}{12} - \frac{t}{8}} \le \exp\left(-\frac{r^2}{2t} + \frac{r^2}{8}\right).$$

Together with (63), this proves the statement for $t \le r$.

To prove the statement in the $0 \le r < t$ case, we will show

$$\int_0^r p_\zeta(u,t)\,du \ge \int_0^r p_Y(u)\,du \quad \text{for } 0 \le r < t \le 1$$

using the lower bound (61) on $p_\zeta(r,t)$. We will prove that for $0 < r < t \le 1$ we have

$$\begin{aligned} p_-(r,t) &= \left(1 + \frac{t}{12}\right)^{-1/2} \frac{1}{t} e^{-\frac{r^2}{2t} - \frac{r^2}{12} - \frac{t}{8}} \sinh(r) \\ &\ge \gamma \sinh\left(\frac{r}{2}\right) \operatorname{sech}^{2\gamma+1}\left(\frac{r}{2}\right) = p_Y(r). \end{aligned}$$

The last inequality is equivalent to

$$e^{-\frac{r^2}{2t} - \frac{r^2}{12}} \ge e^{t/8} \sqrt{1 + \frac{t}{12}} \left(1 - \frac{t}{4}\right) \exp\left(-\left(\frac{4}{t} + 1\right) \log\cosh(r/2)\right). \tag{66}$$

For $0 \le t \le 1$, we have $e^{t/8}\sqrt{1 + \frac{t}{12}}(1 - \frac{t}{4}) \le 1$. Then by the bound (65), we get

$$\begin{aligned} &e^{t/8} \sqrt{1 + \frac{t}{12}} \left(1 - \frac{t}{4}\right) \exp\left(-\left(\frac{4}{t} + 1\right) \log\cosh(r/2)\right) \\ &\quad \le \exp\left(-\left(\frac{4}{t} + 1\right)\left(\frac{r^2}{8} - \frac{r^4}{192}\right)\right). \end{aligned}$$

To complete the proof of (66), we need to show that for $0 < r < t \le 1$ we have

$$\left(\frac{4}{t} + 1\right)\left(\frac{r^2}{8} - \frac{r^4}{192}\right) \ge \frac{r^2}{2t} + \frac{r^2}{12},$$

which follows from direct computation. $\square$

Now we turn to the proof of the total variation bound.

PROOF OF (B) IN LEMMA 9. We need to show that $\int_0^\infty |p_Y(r) - p_\zeta(r,t)|\,dr \le 3t$. By part (a), we have $P(\zeta > r) \le P(Y_\gamma > r)$ which leads to

$$\int_K^\infty \left|p_Y(r) - p_\zeta(r,t)\right| dr \le 2\left(1 - F_Y(K)\right) = 2\operatorname{sech}^{2\gamma}\left(\frac{K}{2}\right), \quad \text{for } K > 0.$$

Setting $K = 2\sqrt{t\log(2/t)} \le 2$, we get

$$\int_K^\infty |p_Y(r) - p_\zeta(r,t)|\,dr \le 2\operatorname{sech}^{2\gamma}\left(\frac{K}{2}\right) \le 2\exp\left(-2\gamma\cdot\frac{5}{12}(K/2)^2\right)$$
$$\le 2(t/2)^{\frac{5}{4}} \le t,$$

where we used (65) in the second inequality.

Using the triangle inequality and the bound (62), we get

$$\int_0^K |p_Y(r) - p_\zeta(r,t)|\,dr \le \int_0^K |p_Y(r) - p_+(r,t)|\,dr + \int_0^K \big(p_+(r,t) - p_\zeta(r,t)\big)\,dr.$$

We can bound the second integral explicitly:

$$\begin{aligned}\int_0^K \big(p_+(r,t) - p_\zeta(r,t)\big)\,dr &\le \int_0^\infty \big(p_+(r,t) - p(r,t)\big)\,dr \\ &= \frac{\sqrt{\frac{\pi}{2}}e^{\frac{3t}{8}}(2\Phi(\sqrt{t})-1)}{\sqrt{t}} - 1 \\ &\le e^{\frac{3t}{8}} - 1 \le t/2.\end{aligned}$$

Introduce

$$p_0(r,t) = \frac{2e^{-\frac{r^2}{2t}}\sinh(\frac{r}{2})}{t}.$$

Then

$$\frac{p_+}{p_0} = e^{-t/8}\cosh(r/2), \qquad \frac{p_Y}{p_0} = (1-t/4)e^{\frac{r^2}{2t}}\operatorname{sech}^{\frac{4}{t}}\left(\frac{r}{2}\right).$$

We have, for $0 \le t \le 1$, $0 \le r \le 2$:

$$\left|\frac{p_+}{p_0} - 1\right| \le \frac{t}{8} + \frac{r^2}{4}.$$

Using the bounds (64) and (65), we get that for $0 \le r \le K < 2$ we have

$$1 = \exp\left(\frac{r^2}{2t} - \frac{4}{t}\frac{r^2}{8}\right) \le e^{\frac{r^2}{2t}}\operatorname{sech}^{\frac{4}{t}}\left(\frac{r}{2}\right) \le \exp\left(\frac{r^2}{2t} - \frac{4}{t}\left(\frac{r^2}{8} - \frac{r^4}{192}\right)\right) = e^{\frac{r^4}{48t}}.$$

This leads to

$$\left|\frac{p_Y}{p_0} - 1\right| \le \frac{t}{4} + |e^{\frac{r^4}{48t}} - 1| \le \frac{t}{4} + \frac{r^4}{24t},$$

where we used that $\frac{r^4}{t} \le t\log(2/t)^2 \le 1$ if $t \le 1$. Note also that $\sinh(x) \le \frac{6}{5}x$ for $x \le 1$. From this,

$$\begin{aligned}\int_0^K |p_Y(r) - p_+(r,t)|\,dr &\le \int_0^K \left(\frac{3}{8}t + \frac{1}{4}r^2 + \frac{r^4}{24t}\right)p_0(r)\,dr \\ &\le \frac{6}{5}\int_0^\infty \left(\frac{3}{8}t + \frac{1}{4}r^2 + \frac{r^4}{24t}\right)\frac{e^{-\frac{r^2}{2t}}r}{t}\,dr = \frac{29}{20}t.\end{aligned}$$

Collecting all our estimates gives

$$\int_0^\infty |p_Y(r) - p_\zeta(r,t)|\,dr \le 3t. \qquad \square$$

The proof of Lemma 11 follows a standard coupling construction.

PROOF OF LEMMA 11. Denote the distributions of X_1, X_2 by μ_1, μ_2, and let $\mu = \frac{1}{2}(\mu_1 + \mu_2)$. Denote that density function of X_i with respect to μ by f_i. From our assumptions, it follows that

$$\tilde{f}_0(x) = \frac{1}{1-\varepsilon}\min\bigl(f_1(x), f_2(x)\bigr), \qquad \tilde{f}_1(x) = \frac{1}{\varepsilon}\bigl|f_1(x) - f_2(x)\bigr|^+,$$
$$\tilde{f}_2(x) = \frac{1}{\varepsilon}\bigl|f_2(x) - f_1(x)\bigr|^+,$$

are also density functions with respect to μ, and the distributions corresponding to $\tilde{f}_1, \tilde{f}_2$ are stochastically ordered just as X_1 and X_2. Moreover, $(1-\varepsilon)\tilde{f}_0 + \varepsilon\tilde{f}_i = f_i$ for $i = 1, 2$.

Recall that if F is a cumulative distribution function, $F^{-1}(x) = \sup\{y : F(y) < x\}$ is its generalized inverse, and U is uniform on $[0, 1]$ then $F^{-1}(U)$ has cumulative distribution function given by F.

Denote the cumulative distribution function corresponding to $\tilde{f}_i$ by $\tilde{F}_i$. Let U_1, U_2 be independent uniform random variables on $[0, 1]$ and consider the pair of random variables

$$\begin{aligned}(Y_1, Y_2) = {}&\mathbf{1}(U_1 \le 1-\varepsilon)\bigl(\tilde{F}_0^{-1}(U_2), \tilde{F}_0^{-1}(U_2)\bigr)\\ &+ \mathbf{1}(U_1 > 1-\varepsilon)\bigl(\tilde{F}_1^{-1}(U_2), \tilde{F}_2^{-1}(U_2)\bigr).\end{aligned}$$

In plain words, with probability $1-\varepsilon$, we generate (Z, Z) where Z has density $\tilde{f}_0$, and with probability ε we generate $(\tilde{X}_1, \tilde{X}_2)$ where $\tilde{X}_i$ has density $\tilde{f}_i$ and $\tilde{X}_1 \le \tilde{X}_2$ a.s. Then $Y_1 \le Y_2$ a.s., $P(Y_1 = Y_2) = 1-\varepsilon$ and Y_i has the same distribution as X_i. Consider the regular conditional distribution of Y_2 given Y_1, and let $g(x, u)$ be the generalized inverse of the conditional cumulative distribution function of Y_2 given $Y_1 = x$. Then $(X_1, g(X_1, U))$ has the same joint distribution as (Y_1, Y_2), and thus g satisfies the requirements of the lemma. $\square$

A.2. Hyperbolic Brownian motion estimates. The first two lemmas give estimates on the behavior of the process $d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$ where $\mathcal{B}$ is a (standard) hyperbolic Brownian motion.

LEMMA 19. *There is a coupling of a hyperbolic Brownian motion $\mathcal{B}$ and a 2-dimensional standard Brownian motion W so that almost surely for all $t \ge 0$ we have*

$$\bigl|W(t)\bigr| \le d_{\mathbb{H}}\bigl(\mathcal{B}(0), \mathcal{B}(t)\bigr) \le \bigl|W(t)\bigr| + t/2.$$

PROOF. The process $q_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$ satisfies the SDE

$$dq = db + \frac{\coth q}{2}\,dt, \qquad q(0) = 0,$$

where b is a standard Brownian motion. This follows, for example, from (10) and the fact that the half-plane representation $x + iy$ of $\mathcal{B}$ satisfies the SDE $d(x + iy) = y(dB_1 + i\,dB_2)$ for $\mathcal{B}$.

Consider the following diffusions with the same driving Brownian motion b as in q:

$$dq_1 = db + \frac{1}{2q_1}\,dt, \qquad dq_2 = db + \left(\frac{1}{2q_2} + \frac{1}{2}\right)dt, \qquad q_1(0) = q_2(0) = 0.$$

Since $\frac{1}{x} \le \coth x \le \frac{1}{x} + 1$ for $x > 0$, we have a.s. $q_1 \le q \le q_2$. (This follows from standard comparison theorems; see, e.g., Ikeda and Watanabe (1989).)The process q_1 is a 2-dimensional Bessel process, and it can be written as the absolute value of a 2-dimensional Brownian motion W.

For the upper bound, note that we have $q_1 \le q_2$, and taking the difference of the SDEs for q_2 and q_1 we get

$$(q_2 - q_1)' = \frac{1}{2q_2} - \frac{1}{2q_1} + \frac{1}{2} \le \frac{1}{2}.$$

Thus $q(t) \le q_2(t) \le q_1(t) + t/2 = |W(t)| + t/2$, which completes the proof. $\square$

LEMMA 20. *Let $\mathcal{B}$ be a standard hyperbolic Brownian motion. Then for $t > 0, a > 0$ we have*

$$P\Big(\max_{0\le s\le t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \le a\Big) \le \frac{4}{\pi} e^{-\frac{\pi^2 t}{8a^2}}. \tag{67}$$

If $0 < t \le a$, then

$$P\Big(\max_{0\le s\le t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \ge a\Big) \le \frac{16\sqrt{t}}{a\sqrt{\pi}} e^{-\frac{a^2}{16t}}. \tag{68}$$

PROOF. By Lemma 19, the process $q_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$ stochastically dominates $|W(t)|$ where $W = (B_1, B_2)$ is a 2-dimensional Brownian motion. Then

$$\begin{aligned} P\Big(\max_{0\le s\le t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \le a\Big) &\le P\Big(\max_{0\le s\le t} |W(s)| \le a\Big) \\ &\le P\Big(\max_{0\le s\le t} |B_1(s)| \le a\Big) \le \frac{4}{\pi} e^{-\frac{\pi^2 t}{8a^2}}, \end{aligned}$$

which proves (67). The last step follows from the following identity for the standard Brownian motion B_1 (see, e.g., Section 7.4 of Mörters and Peres (2010)):

$$P\Big(\max_{0\le s\le t} |B_1(s)| \le u\Big) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} e^{-\frac{(2k+1)^2\pi^2 t}{8u^2}}.$$

From Lemma 19, it follows that $q_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$ is stochastically dominated by the process $|W(t)| + t/2$ where W is a 2-dimensional standard Brownian motion. Thus

$$\begin{aligned} &P\Big(\max_{0\le s\le t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \ge a\Big) \\ &\quad\le P\Big(\max_{0\le s\le t} \big(|W(s)| + s/2\big) \ge a\Big) \le P\Big(\max_{0\le s\le t} |W(s)| \ge a - t/2\Big) \\ &\quad\le 4P\Big(\max_{0\le s\le t} B(s) \ge \frac{1}{\sqrt{2}}(a - t/2)\Big) = 4P\Big(|B(t)| \ge \frac{1}{\sqrt{2}}(a - t/2)\Big) \\ &\quad\le \frac{8\sqrt{t}}{(a - t/2)\sqrt{\pi}} e^{-\frac{(a-t/2)^2}{4t}} \le \frac{16\sqrt{t}}{a\sqrt{\pi}} e^{-\frac{a^2}{16t}}. \end{aligned}$$

In the second line, we used that $W(s) \ge x$ implies that one of the coordinates of W is at least $\frac{1}{\sqrt{2}}x$ in absolute value, then used the reflection principle. Finally, we used the well-known tail bound for the normal distribution and $t \le a$. $\square$

The next lemma shows that $\mathcal{B}$ approaches its limit point with speed $1/2$. (Note that proved bound is not optimal.)

LEMMA 21. *Let $\mathcal{B}$ be a hyperbolic Brownian motion and let $z(t)$ be the point moving with speed $1/2$ on the geodesic connecting $\mathcal{B}(0)$ to $\mathcal{B}(\infty) = \lim_{t\to\infty} \mathcal{B}(t)$. Then there is a random $C < \infty$ so that almost surely*

$$d_{\mathbb{H}}\big(\mathcal{B}(t), z(t)\big) \le C + t^{1/2} \log(1+t)$$

for all $t \ge 0$.

PROOF. Consider the half-plane representation of $\mathbb{H}$ where $\mathcal{B}(0) = i$ and $\mathcal{B}(\infty) = \infty$, and denote the representation of $\mathcal{B}$ by $x + iy$. Then $x + iy$ is a hyperbolic Brownian motion conditioned to hit ∞, in particular, it satisfies

$$dy = y(dB_1 + dt), \qquad dx = y\, dB_2, \qquad y(0) = 1, \qquad x(0) = 0,$$

where B_1, B_2 are independent standard Brownian motions. (See, e.g., Valkó and Virág (2017).) The geodesic connecting i and ∞ is $\{ie^t, t \ge 0\}$, and the point moving with speed $1/2$ is $z(t) = ie^{t/2}$. By the triangle inequality,

$$\begin{aligned} d_{\mathbb{H}}\big(\mathcal{B}(t), z(t)\big) &\le d_{\mathbb{H}}(x+iy, iy) + d_{\mathbb{H}}\big(iy, z(t)\big) \\ &= \operatorname{arccosh}\Big(1 + \frac{x^2}{2y^2}(t)\Big) + \big|\log\big(y(t)e^{-t/2}\big)\big| \\ &\le \log\Big(2 + \frac{x^2}{y^2}(t)\Big) + \big|\log\big(y(t)e^{-t/2}\big)\big|. \end{aligned} \tag{69}$$

We can explicitly solve for y and x from the SDE:

$$y(t) = e^{B_1(t)+t/2}, \qquad x(t) = \int_0^t e^{B_1(s)+s/2}\, dB_2(s).$$

We have $|\log(y(t)e^{-t/2})| = |B_1(t)|$, and using the law of iterated logarithm we get the bound

$$\big|\log\big(ye^{-t/2}\big)\big| = \big|B_1(t)\big| \le C_0 + \sqrt{t\log(1+t)} \tag{70}$$

for all $t \ge 0$ with some random C_0 depending on B_1.

To bound $\log(2 + \frac{x^2}{y^2})$, we start with the observation that there is a standard Brownian motion B so that $x(t) = B(\int_0^t y(s)^2\, ds)$. Thus, using the law of iterated logarithm again,

$$\begin{aligned} \Big(\frac{x(t)}{y(t)}\Big)^2 &\le y(t)^{-2}\Big(2C_0^2 + 2\int_0^t y(s)^2\, ds \log\Big(1 + \int_0^t y(s)^2\, ds\Big)\Big) \\ &= 2C_0 e^{-2B(t)-t} + 2\int_0^t e^{2(B(s)-B(t))-t+s}\, ds \log\Big(1 + \int_0^t e^{2B(s)+s}\, ds\Big). \end{aligned}$$

Using the bound $|B(s)| \le C_0 + \sqrt{t\log(1+t)}$ for $0 \le s \le t$ leads to

$$\Big(\frac{x(t)}{y(t)}\Big)^2 \le C_1\big(1 + te^{4\sqrt{t\log(1+t)}}\big)$$

with a random C_1 depending only on $\mathcal{B}(\cdot)$. Using (69) and (70), the statement follows. □

The next statement gives an estimate on the modulus of continuity of the hyperbolic Brownian motion. The proof follows that of the analogous statement for standard Brownian motion, we include it for completeness. The constants have not been optimized.

PROPOSITION 22. *Let $\mathcal{B}$ be a hyperbolic Brownian motion. Then there is a random constant* $0 < h_0 \le 1$ *so that a.s.*,

$$d_{\mathbb{H}}\big(\mathcal{B}(s), \mathcal{B}(s+h)\big) \le 20\sqrt{h \log\Big(2 + \frac{s+1}{h}\Big)}, \tag{71}$$

for all $0 < h \le h_0$ *and* $0 \le s$.

Note that the proof below also shows that there is a random constant C so that (71) holds for all $0 < h \le 1$ with C in place of 20.

PROOF. Let $I_{m,n} = [m2^{-n}, (m+1)2^{-n}]$ and $\Delta_{m,n} = \max_{t \in I_{m,n}} d_{\mathbb{H}}(\mathcal{B}(m2^{-n}), \mathcal{B}(t))$ for $m, n \ge 0$. If $2^{-n/2} < u$, then by Lemma 20 we have

$$P\big(\Delta_{m,n} \ge 2^{-n/2}u\big) = P\Big(\max_{0 \le t \le 2^{-n}} d_{\mathbb{H}}\big(\mathcal{B}(0), \mathcal{B}(t)\big) \ge 2^{-n/2}u\Big) \le \frac{16}{u\sqrt{\pi}} e^{-\frac{u^2}{16}}.$$

Thus for $n \ge 0$, $m \ge 0$ we get

$$P\bigg(\Delta_{m,n} \ge \frac{9}{2} \cdot 2^{-n/2}\sqrt{\log(2^n + m + 1)}\bigg) \le \frac{3(2^n + m + 1)^{-5/4}}{\sqrt{\log(2^n + m + 1)}}.$$

We have $\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} (2^n + m + 1)^{-5/4} < \infty$. By the Borel–Cantelli lemma, there is a random $N_0 \ge 1$ so that if $n \ge N_0$ then

$$\Delta_{m,n} \le \frac{9}{2}\sqrt{2^{-n} \log(2^n + m + 1)}. \tag{72}$$

We will show that (71) holds with $0 < h \le h_0 = 2^{-N_0}$ and $0 \le s$. Let m, n be nonnegative integers with $2^{-n-1} < h \le 2^{-n}$ and $m2^{-n} \le s \le (m+1)2^{-n}$. Then we have $n \ge N_0$, and using the triangle inequality and (72) we get

$$\begin{aligned} d_{\mathbb{H}}\big(\mathcal{B}(s), \mathcal{B}(s+h)\big) &\le 2\Delta_{m,n} + \Delta_{m+1,n} \le 3 \cdot \frac{9}{2}\sqrt{2^{-n} \log(2^n + m + 2))} \\ &\le 20\sqrt{h \log\Big(2 + \frac{s+1}{h}\Big)}, \end{aligned}$$

which completes the proof. $\square$

Our next proposition gives a lower bound on the fluctuations of the hyperbolic Brownian motion. The proof again follows that of the analogous statement for standard Brownian motion.

PROPOSITION 23. *Let $\mathcal{B}$ be a hyperbolic Brownian motion. Then there is a random constant* $C_0 < \infty$ *so that a.s. for any* $0 \le s < t$ *there exists* $u, v \in [s, t]$ *with*

$$d_{\mathbb{H}}\big(\mathcal{B}(u), \mathcal{B}(v)\big)^2 \ge C_0 \frac{t - s}{\log(2 + t) + \log(t - s + \frac{1}{t-s})}. \tag{73}$$

Moreover, with the same constant C_0 *a.s. for any* $0 \le s < t$ *there exists* $u \in [s, t]$ *with*

$$d_{\mathbb{H}}\big(\mathcal{B}(s), \mathcal{B}(v)\big)^2 \ge \frac{1}{4} C_0 \frac{t - s}{\log(2 + t) + \log(t - s + \frac{1}{t-s})}. \tag{74}$$

PROOF. The second part of the statement follows from the first part using the triangle inequality.

To prove (73), it is enough to show the statement for pairs of the form $s = m2^n$, $t = (m+1)2^n$ with $m, n \in \mathbb{Z}$ and $m > 0$. We partition the interval $[s, t)$ into $k = \lceil 1+|n|+5\log(2+m)\rceil$ subintervals $[a_i, a_{i+1})$ of size $\frac{2^n}{k}$. For a given $0 \le i < k$, we have

$$P\left(d_{\mathbb{H}}(\mathcal{B}(a_i), \mathcal{B}(a_{i+1})) \le \frac{2^{n/2}}{\sqrt{k}}\right) = P\left(d_{\mathbb{H}}\left(\mathcal{B}(0), \mathcal{B}\left(\frac{2^n}{k}\right)\right) \le \frac{2^{n/2}}{\sqrt{k}}\right)$$
$$\le P\left(\left|B\left(\frac{2^n}{k}\right)\right| \le \frac{2^{n/2}}{\sqrt{k}}\right) \le P(|B(1)| \le 1) \le \frac{4}{5},$$

where $B(t)$ is a standard Brownian motion and we used Lemma 19. Using the Markov property of the hyperbolic Brownian motion, we get that

$$P\left(d_{\mathbb{H}}(\mathcal{B}(a_i), \mathcal{B}(a_{i+1})) \le \frac{2^{n/2}}{\sqrt{k}} \text{ for all } 0 \le i < k\right) \le \left(\frac{4}{5}\right)^k \le \left(\frac{4}{5}\right)^{|n|+5\log(2+m)}.$$

Since $\sum_{m=1}^{\infty}\sum_{n\in\mathbb{Z}}(\frac{4}{5})^{|n|+5\log(2+m)} < \infty$, there are a.s. finitely many pairs m, n for which we cannot find $u, v \in [m2^n, (m+1)2^n]$ with

$$d_{\mathbb{H}}(\mathcal{B}(u), \mathcal{B}(v))^2 \le \frac{2^n}{5\log(2+m)+|n|}.$$

Now (73) follows with a random C_0 for pairs of the form $s = m2^n$, $t = (m+1)2^n$, and from this (73) follows with a modified C_0 for all s, t. □

Acknowledgments. We thank Paul Bourgade, Yun Li and Bálint Vető for interesting discussions and comments. We thank the anonymous referee for valuable comments and suggestions, especially for suggesting to investigate the $\beta' = \infty$ case in Theorem 3.

The first author was supported in part by NSF CAREER award DMS-1053280, NSF award DMS-1712551 and the Simons Foundation.

The second author was supported by the Canada Research Chair program the NSERC Discovery Accelerator grant, the MTA Momentum Random Spectra research group and the ERC consolidated grant 648017 (Abért).

REFERENCES

ALLEZ, R. and DUMAZ, L. (2014). From sine kernel to Poisson statistics. *Electron. J. Probab.* **19** no. 114, 25. MR3296530 https://doi.org/10.1214/EJP.v19-3742

ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

BHATIA, R. and ELSNER, L. (1994). The Hoffman–Wielandt inequality in infinite dimensions. *Proc. Indian Acad. Sci. Math. Sci.* **104** 483–494. MR1314392 https://doi.org/10.1007/BF02867116

BOURGADE, P., NAJNUDEL, J. and NIKEGHBALI, A. (2013). A unitary extension of virtual permutations. *Int. Math. Res. Not. IMRN* **18** 4101–4134. MR3106884 https://doi.org/10.1093/imrn/rns167

DAVIES, E. B. and MANDOUVALOS, N. (1988). Heat kernel bounds on hyperbolic space and Kleinian groups. *Proc. Lond. Math. Soc.* (3) **57** 182–208. MR0940434 https://doi.org/10.1112/plms/s3-57.1.182

DUMITRIU, I. and EDELMAN, A. (2002). Matrix models for beta ensembles. *J. Math. Phys.* **43** 5830–5847. MR1936554 https://doi.org/10.1063/1.1507823

EDELMAN, A. and SUTTON, B. D. (2007). From random matrices to stochastic operators. *J. Stat. Phys.* **127** 1121–1165. MR2331033 https://doi.org/10.1007/s10955-006-9226-4

FORRESTER, P. J. (2010). *Log-Gases and Random Matrices*. *London Mathematical Society Monographs Series* **34**. Princeton Univ. Press, Princeton, NJ. MR2641363 https://doi.org/10.1515/9781400835416

HOLCOMB, D. and VALKÓ, B. (2015). Large deviations for the Sine_β and Sch_τ processes. *Probab. Theory Related Fields* **163** 339–378. MR3405620 https://doi.org/10.1007/s00440-014-0594-4

IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. MR1011252

KARPELEVIČ, F. I., TUTUBALIN, V. N. and ŠUR, M. G. (1959). Limit theorems for compositions of distributions in the Lobačevskiĭ plane and space. *Theory Probab. Appl.* **4** 399–402.

KILLIP, R. and NENCIU, I. (2004). Matrix models for circular ensembles. *Int. Math. Res. Not.* **50** 2665–2701. MR2127367 https://doi.org/10.1155/S1073792804141597

KILLIP, R. and STOICIU, M. (2009). Eigenvalue statistics for CMV matrices: From Poisson to clock via random matrix ensembles. *Duke Math. J.* **146** 361–399. MR2484278 https://doi.org/10.1215/00127094-2009-001

MAPLES, K., NAJNUDEL, J. and NIKEGHBALI, A. (2019). Strong convergence of eigenangles and eigenvectors for the circular unitary ensemble. *Ann. Probab.* **47** 2417–2458. MR3980924 https://doi.org/10.1214/18-AOP1311

MECKES, E. S. and MECKES, M. W. (2016). Self-similarity in the circular unitary ensemble. *Discrete Anal.* Paper No. 9, 14. MR3555192 https://doi.org/10.19086/da.736

MEHTA, M. L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics* (*Amsterdam*) **142**. Elsevier/Academic Press, Amsterdam. MR2129906

MÖRTERS, P. and PERES, Y. (2010). *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics* **30**. Cambridge Univ. Press, Cambridge. MR2604525 https://doi.org/10.1017/CBO9780511750489

NAKANO, F. (2014). Level statistics for one-dimensional Schrödinger operators and Gaussian beta ensemble. *J. Stat. Phys.* **156** 66–93. MR3215116 https://doi.org/10.1007/s10955-014-0987-x

RAMÍREZ, J. A. and RIDER, B. (2009). Diffusion at the random matrix hard edge. *Comm. Math. Phys.* **288** 887–906. MR2504858 https://doi.org/10.1007/s00220-008-0712-1

RAMÍREZ, J. A., RIDER, B. and VIRÁG, B. (2011). Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.* **24** 919–944. MR2813333 https://doi.org/10.1090/S0894-0347-2011-00703-0

VALKÓ, B. and VIRÁG, B. (2009). Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177** 463–508. MR2534097 https://doi.org/10.1007/s00222-009-0180-z

VALKÓ, B. and VIRÁG, B. (2017). The Sine_β operator. *Invent. Math.* **209** 275–327. MR3660310 https://doi.org/10.1007/s00222-016-0709-x

WIGNER, E. P. (1951). On a class of analytic functions from the quantum theory of collisions. *Ann. of Math.* (2) **53** 36–67. MR0039059 https://doi.org/10.2307/1969342